CAMILA ISOTON

CONDIÇÕES NECESSÁRIAS E SUFICIENTES DE OTIMALIDADE PARA PROBLEMAS COM UM E COM VÁRIOS OBJETIVOS

CAMILA ISOTON

CONDIÇÕES NECESSÁRIAS E SUFICIENTES DE OTIMALIDADE PARA PROBLEMAS COM UM E COM VÁRIOS OBJETIVOS

Dissertação apresentada ao Programa de Pós-Graduação em Matemática Aplicada da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Mestre em Matemática Aplicada.
Orientadora: Prof.^{a} Dr.^{a} Lucelina Batista dos Santos.

Curitiba
Agosto de 2013

UNIVERSIDADE FEDERAL DO PARANÁ

CAMILA ISOTON

CONDIÇÕES NECESSÁRIAS E SUFICIENTES DE OTIMALIDADE PARA PROBLEMAS COM UM E COM VÁRIOS OBJETIVOS

Prof^{a}. Dr^{a}. Lucelina Batista dos Santos
Orientadora - Universidade Federal do Paraná - UFPR

Prof. Dr. Marko Antonio Rojas Medar
Universidad del Bio-Bio - Chile

Prof. Dr. Lucas Garcia Pedroso
Universidade Federal do Paraná - PR

Curitiba, 23 de agosto de 2013.

Agradecimentos

Meus sinceros agradecimentos:

Agradeço primeiramente a minha querida orientadora, pela paciência, amizade e ajuda. Pelo tempo, esforço e dedicação empregados nesta pesquisa.

Aos meus pais Expedito e Lorena, e a minha irmã, Gisele, que estiveram sempre presentes ao meu lado (mesmo longe fisicamente), me apoiando, incentivando, ajudando, compreendendo, enfim, por todo o carinho e amor que sempre tiveram comigo. Aos meus familiares que tanto me apoiaram e me incentivaram.

Agradeço aos meus queridos colegas do PPGMA, Ana, Stela e Geovani, ao Dion pela ajuda, e a todos os demais pelos conhecimentos compartilhados, amizade e companheirismo. Este trabalho tem um pouco de cada um de vocês.

Aos professores do programa, que me ajudaram a crescer intelectual e pessoalmente, principalmente aos professores Ademir e Elizabeth, pelo apoio e incentivo. Aos ex-colegas e professores do Departamento de Matemática da UTFPR-Campus Pato Branco, principalmente Adriano e Gilberto pela ajuda e incentivo. Aos membros da banca, professores Marko, Washington e Lucas, obrigada por aceitarem o convite.

As amigas Cristina e Patrícia, pelo apoio, compreensão e amizade ao longo destes dois anos. Aos demais amigos que me apoiaram. Muito Obrigada.

Ao Programa de Pós-Graduação em Matemática Aplicada da UFPR pela oportunidade e formação de qualidade propiciada. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPQ, pelo apoio financeiro.

E finalmente, mas não menos importante, a Deus pela capacitação concedida, pela força espiritual e por seguir ao meu lado.

“ O que é o homem na natureza? Um nada em relação ao infinito, um tudo em relação ao nada, um ponto a meio entre nada e tudo”

Blaise Pascal

Aos meus amados pais
Expedito e Lorena, por sempre confiarem, acreditarem, apoiarem e me darem forças para seguir.

Resumo

Apresenta-se um estudo da existência de soluções e das condições de otimalidade de primeira e segunda ordem, para problemas com um e com vários objetivos. Neste estudo, a convexidade desempenha um papel muito importante. Nosso objetivo foi estender os conceitos de convexidade para tais problemas, nos quais as funções envolvidas são diferenciáveis. Definimos os conceitos de pontos estacionários e de condições de Kuhn-Tucker para problemas mono-objetivos e os seus análogos para os problemas multiobjetivos. Neste âmbito, os Teoremas da Alternativa foram largamente empregados. Introduzir-se-ão conceitos de invexidade, nos quais, os problemas KT-invexos e KT-pseudoinvexos foram a chave para garantir a suficiência destas condições.

Palavras-chave: Condições de Otimalidade; otimização mono e multiobjetivos; funções convexas generalizadas, KT-invexidade.

Abstract

In this work we study the existence of the solutions and the first- and second-order optimality conditions (ou first and second order optimality conditions) for scalar and vector optimization problems. In this study, the plays a very important role. The main aspect of this work was to extend the concepts of convexity for such optimization problems, where the functions are differentiable. We define concepts of stationary points and Kuhn-Tucker conditions for mono-objective problems and for their analogues multi-objective problems. In this context, Theorems of Alternative have been widely used. Invexity concepts have been used, in which the problems KT-invexos and KT-pseudoinvexos was key to ensure the sufficiency of these conditions.

Keywords: Optimality Conditions; scalar and vector (ou mono-objective and multi-objective) optimization problems; generalized convexity, KT-invexity.

Sumário

Capítulo 1

Introdução

Diariamente encontramos situações que nos levam à tomada de decisões. Esse tipo de problema surge por exemplo, nas áreas da Engenharia, Economia, Administração, etc. Inicialmente, restrito a problemas que envolvem um único objetivo, os modelos de otimização e suas técnicas visam minimizar (ou maximizar) uma função (nem sempre linear), denominada função objetivo, de várias variáveis escolhidas dentro de alguma região admissível (factível), que é definida pelas restrições impostas às variáveis do problema.

Esses problemas que buscam calcular o ponto ótimo (de mínimo ou máximo) de alguma função objetivo, sujeita ou não a restrições, são chamados problemas de programação matemática.

No sentido de contemplar situações mais realistas, os problemas de programação matemática evoluíram para o estudo de vários objetivos, em geral conflitantes, que concorrem para uma determinada solução. Esses problemas de otimização com objetivos múltiplos, denotados problemas multiobjetivos, são interessantes tanto do ponto de vista teórico quanto aplicado a problemas reais (para aplicações veja, por exemplo, [26] e admitem a seguinte formulação:

$$\begin{array}{l} \text{minimizar } f(x) = (f_1(x), \ldots, f_p(x)) \\ \text{sujeito a } x \in D \end{array} \tag{1.1}$$

em que $D \subset \mathbb{R}^n$ é um conjunto não vazio e $f : \mathbb{R}^n \to \mathbb{R}$.

Conceitualmente, um ponto ótimo de (1.1) é chamado eficiente, quando não é possível melhorar nenhum objetivo sem piorar em algum outro. Essa nomenclatura é devido a Pareto, um dos precursores neste campo da Otimização Multiobjetivo, que tratou do tema em seu célebre trabalho *"Cours d'Econimie Politique"* [30].

Em meados da década de 40 muitos trabalhos surgiram no campo da otimização (auge da Pesquisa Operacional), os quais se destacam Charnes e Cooper, Gale, Kuhn e Tucker e de Karlin. Entre 60 e 70 a literatura sobre este tema foi se sofisticando e noções como *eficiência fraca* e *eficiência própria* surgiram. A grosso modo, um ponto é fracamente eficiente se não é possível melhorar todos os objetivos simultaneamente e propriamente eficiente se é solução eficiente e se os quocientes entre ganho em um objetivo e perda com respeito aos demais é limitada.

A partir disso, inúmeras publicações têm aparecido visando desenvolver critérios que permitam decidir quando um ponto factível é ou não eficiente. Um dos métodos mais clássicos tenta relacionar o problema multiobjetivo com algum problema escalar específico, cuja teoria é bastante desenvolvida.

Nesse âmbito, a convexidade desempenha um papel fundamental. Sob hipóteses de convexidade as condições necessárias e suficientes de otimalidade são trivialmente estabelecidas, as quais garantem a existência de pontos ótimos (eficientes).

Desde a última década outras famílias de funções mais gerais que as funções convexas começaram a surgir no intuito de generalizar o conceito de convexidade. Tais funções são conhecidas por funções convexas generalizadas. Entre elas, as funções invexas introduzidas por Hanson [16] tem grande destaque, uma vez que esta classe de funções tem a propriedade de que todo ponto é mínimo global quando, e somente quando, for estacionário. Além disso, para problemas definidos por restrições de desigualdades, as condições de Kuhn-Tucker são suficientes para a otimalidade, quando as funções envolvidas no problema são invexas. Além disso, Martin [4], introduziu uma classe de problemas, denominados Kuhn-Tucker invexos (KT-invexos), que inclui a classe dos problemas invexos, a qual satisfaz a propriedade: o problema é KT-invexo quando, e somente quando, todo ponto que satisfaz as condições de Kuhn-Tucker é minimo global. Contrapartidas deste resultado para o problema vetorial (1.1) foram estabelecidas em Osuna-Gómez et al.[25] para soluções fracamente eficientes.

Até agora, quando citamos as condições necessárias e suficientes de otimalidade estavamos na verdade nos referindo as condições de otimalidade de primeira ordem. É sabido que as condições de otimalidade de segunda ordem são de suma importância, já que elas nos auxiliam a eliminar candidatos não ótimos em nossos problemas de programação matemática. As condições necessárias de segunda ordem são utilizadas para eliminar pontos estacionários não ótimos, e as suficientes de segunda ordem nos ajudam a determinar quando um dado ponto é de fato minimizador.

Trabalhando com convexidade generalizada, à luz das condições de segunda ordem, Ginchev e Ivanov [27] fizeram para problemas com um objetivo um interessante trabalho usando a noção de segunda ordem e de derivadas direcionais. Além disso, Ivanov [28] introduziu a noção de KT-invexidade de segunda ordem e provou que um problema é KT-invexo de segunda ordem quando, e somente quando, todo ponto que satisfaz condições necessárias de segunda ordem é um minimizador global deste problema.

Visando estender os conceitos explorados por Ginchev e Ivanov [27, 28] aos problemas multiobjetivos, Santos, et al. [15] definem pontos de Kuhn-Tucker de segunda ordem e deriva as condições necessárias de otimalidade de segunda ordem para problemas multiobjetivos. Além disso, introduzem o conceito de KT-pseudoinvexidade de segunda ordem cuja principal propriedade é: o problema multi-objetivo é KT-pseudoinvexo de segunda ordem quando, e somente quando, todo ponto Kuhn-Tucker de segunda ordem do problema é solução fracamente eficiente.

Este trabalho divide-se em duas partes. A primeira, concentra-se no estudo de problemas de programação matemática com um objetivo. O intuito é estudar a existência de soluções para tais problemas e garantir as condições necessárias e suficientes de otimalidade. A segunda, trata dos problemas de programação matemática multiobjetivos, e uma análise análoga a realizada na primeira parte é estudada.

Parte I

Problema Mono-objetivo

Capítulo 2

Existência de Soluções

Inicialmente, consideraremos o seguinte problema de otimização:

$$\begin{array}{l} \text{Minimizar} \;\; f(x) \\ \text{sujeito a } x \in D \end{array} \tag{2.1}$$

onde $D \subset \mathbb{R}^n$ é um conjunto não vazio e $f : \mathbb{R}^n \to \mathbb{R}$.

Para este problema, a função f será chamada função objetivo e D é o conjunto de restrições (também chamado conjunto viável). Cada ponto $x \in D$ será chamado ponto factível (ou ponto viável).

Os conceitos de solução para o problema (2.1) são os seguintes:

- Um ponto factível $x^* \in D$ é um *minimizador* (ou solução) do problema (2.1) se

$$f(x^*) \leq f(x), \forall x \in D.$$

- Um ponto factível $x^* \in D$ é um *minimizador local* do problema (2.1) se existe uma vizinhança V do ponto x^* tal que

$$f(x^*) \leq f(x), \forall x \in D \cap V.$$

- O *valor ótimo do problema* (2.1), se define com $v = \inf_{x \in D} f(x)$.

A figura a seguir ilustra os conceitos de minimizador e de minimizador local.

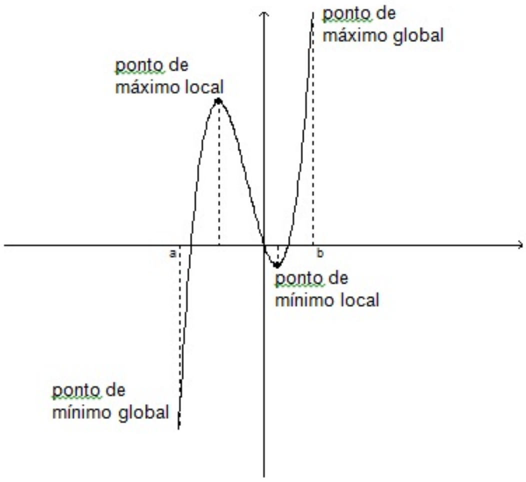


Figura 2.1: Caracterização de máximos e mínimos

Na primeira parte desta Dissertação consideraremos o seguinte problema de otimização:

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a:} & \begin{array}{l} g_i(x) \leq 0, i = 1, \cdots, m \\ x \in U \end{array} \end{array} \tag{P}$$

onde $f, g_i : \mathbb{R}^n \to \mathbb{R}$ e $U \subset \mathbb{R}^n$ é um subconjunto aberto e não vazio de $\mathbb{R}^n$.

Por simplicidade, definiremos $I = \{1, \cdots, m\}$.

Denotaremos

$$X = \{x \in U : g_i(x) \leq 0, i \in I\}$$

ao conjunto factível. Note que, o problema (P) é um caso particular do problema (2.1), tomando-se $D = X$.

Cada função $g_i : \mathbb{R}^n \to \mathbb{R}$ é uma restrição (de desigualdade) do problema (P).

Para cada ponto factível $x \in X$, definiremos o conjunto

$$I(x) = \{i \in I : g_i(x) = 0\}$$

o qual é chamado conjunto dos índices das restrições ativas em x.

O objetivo fundamental deste capítulo é estudar os resultados de existência de soluções para o problema (P). Antes porém, recordamos o muito conhecido Teorema de Weirstrass, o qual está muito relacionado a este tema:

Proposição 2.1 *[Teorema de Weierstrass] Suponha que D é um subconjunto compacto e não-vazio de $\mathbb{R}^n$ e que a função $f : \mathbb{R}^n \to \mathbb{R}$ é contínua em D. Então existem $x_1, x_2 \in D$ tais que*

$$f(x_1) \leq f(x) \leq f(x_2), \forall x \in D.$$

Para demonstração, veja [6].

Em particular, se f é contínua e o conjunto factível D é compacto e não vazio, então o problema (P) tem solução.

Neste capítulo estudaremos alguns resultados de existência de soluções para o problema (P) com hipóteses mais fracas de continuidade (na função objetivo f) e de compacidade do conjunto factível X.

Para isto, iniciamos recordando as definições de semi-continuidade:

Definição 2.2 *Seja a função* $f : \mathbb{R}^n \to \mathbb{R}$.

(1) Dizemos que f *é* ***semi-contínua inferiormente*** *(sci, por simplicidade)* ***em*** $\overline{x} \in \mathbb{R}^n$, *se para cada* $\varepsilon > 0$ *existe um* $r = r(\varepsilon) > 0$ *tal que*

$$f(x) > f(\overline{x}) - \varepsilon, \forall x \in B(\overline{x}, r)$$

(onde $B(\overline{x}, r)$ *denota a bola aberta de centro* $\overline{x}$ *e raio* r*). Se* f *for sci em cada* $x \in \mathbb{R}^n$ *diremos simplesmente que* f *é semi-contínua inferiormente.*

(2) Dizemos que f *é* ***semi-contínua superiormente*** *(scs, por simplicidade)* ***em*** $\overline{x}$ *se*$(-f)$ *for sci em* $\overline{x}$*. Similarmente, se* f *for scs em cada* $x \in \mathbb{R}^n$ *diremos simplesmente que* f *é semi-contínua superiormente.*

Note que a noção de semi-continuidade é mais fraca que a de continuidade. Se f é contínua em $\overline{x}$, então f é sci e scs em $\overline{x}$. O exemplo a seguir mostra que semi-continuidade inferior não implica continuidade.

Exemplo 2.3 *Seja a função* $f : \mathbb{R} \to \mathbb{R}$ *definida por:*

$$f(x) = \begin{cases} x^2, & se\ x \leq 1 \\ -2x^2 + 4x, & se\ x > 1. \end{cases} \tag{2.2}$$

Temos que f *é semi-contínua inferior em* $x = 1$*, entretanto, pelo gráfico (Figura 2.2) abaixo, podemos observar que o limite lateral inferior tende a 1 e o superior a 2, logo* f *não é contínua em* $x = 1$.

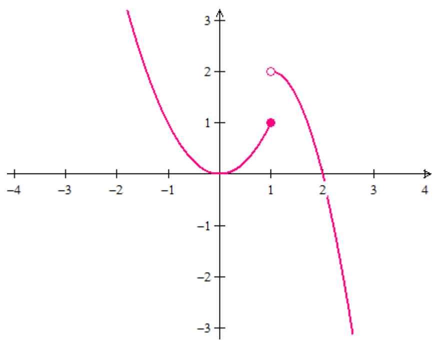


Figura 2.2: Função semi-contínua inferior

A seguinte proposição nos dá uma caracterização alternativa de semi-continuidade:

Proposição 2.4 *Uma função* $f : \mathbb{R}^n \to \mathbb{R}$ *é sci em* $\overline{x} \in \mathbb{R}^n$ *se, e somente se, para toda sequencia* (x^k) *convergente a* $\overline{x}$*, tem-se*

$$f(\overline{x}) \leq \liminf_{k\to\infty} f(x^k). \tag{2.3}$$

Demonstração. Seja f uma função sci em $\overline{x}$ e considere uma sequencia (x^k), com $x^k \to \overline{x}$. Tome $\varepsilon > 0$ arbitrário. Sendo f sci em $\overline{x}$, existe uma bola aberta $B(\overline{x}, r)$, tal que $f(x) > f(\overline{x}) - \varepsilon, \forall x \in B(\overline{x}, r)$. Como $x^k \to \overline{x}$, existe um $n_0 \in \mathbb{N}$ tal que $x^k \in B(\overline{x}, r), \forall k > n_0$. Em particular, $f(x^k) > f(\overline{x}) - \varepsilon, \forall k > n_0$. Em particular,

$$\liminf_{k\to\infty} f(x^k) \geq f(\overline{x}) - \varepsilon$$

e, sendo $\varepsilon > 0$ arbitrário, $\liminf_{k\to\infty} f(x^k) \geq f(\overline{x})$.

Para demonstração da recíproca, suponha por contradição que (2.3) é válido para cada sequência $x^k \to \overline{x}$ e que f não é sci em $\overline{x}$. Então, existe um $\varepsilon' > 0$, tal que, para cada $r > 0$, existe um $x \in B(\overline{x}, r)$ tal que $f(x) \leq f(\overline{x}) - \varepsilon'$. Assim sendo, existe uma sequência $z^k \to \overline{x}$ tal que $f(z^k) \leq f(\overline{x}) - \varepsilon', \forall k \in \mathbb{N}$. Da hipótese feita,

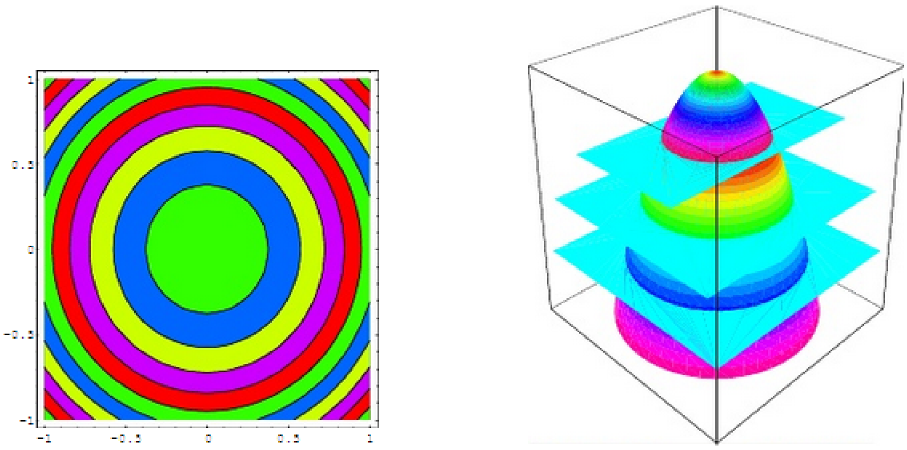
$$f(\overline{x}) \leq \liminf_{k\to\infty} f(z^k) \leq f(\overline{x}) - \varepsilon'$$

o que é absurdo. $\square$

Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função dada e $\alpha \in \mathbb{R}$.

Se define o *conjunto de nível*,

$$S_\alpha(f) = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}.$$



(a) (b)

Figura 2.3: Conjunto de Níveis

Observe os itens (a) e (b) da Figura 2.3. Em (a) está ilustrado o gráfico das curvas de nível da função representada em (b), onde os planos cortantes estão representados em azul claro. Note que curvas de nível muito espaçadas significa que o gráfico cresce lentamente; duas curvas de nível muito próximas significa que o gráfico cresce abruptamente.

O seguinte resultado nos dá uma caracterização de sci em termos dos conjuntos de nível.

Proposição 2.5 *Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é sci se, e somente, se $S_\alpha(f)$ é fechado, para cada $\alpha \in \mathbb{R}$.*

Demonstração. Sejam f uma função sci e $\alpha \in \mathbb{R}$. Mostraremos que o conjunto $S_\alpha(f)$ é fechado. Para isto, provaremos que seu complementar, $[S_\alpha(f)]^c$ é aberto. Tome $\overline{x} \in [S_\alpha(f)]^c$, então, $f(\overline{x}) > \alpha$ e, para cada $\varepsilon > 0$, existe uma bola $B(\overline{x}, r)$, tal que, $f(x) > f(\overline{x}) - \varepsilon, \forall x \in B(\overline{x}, r)$. Então, $f(x) > \alpha - \varepsilon$ e, como $\varepsilon > 0$ é arbitrário, temos que $f(x) > \alpha$ para cada $x \in B(\overline{x}, r)$. Ou seja, $[S_\alpha(f)]^c$ é aberto.

Agora, demonstramos a recíproca. Suponha que o conjunto $[S_\alpha(f)]^c$ é aberto, para todo $\alpha \in \mathbb{R}$. Fixe $\overline{x} \in \mathbb{R}^n$ e $\varepsilon > 0$. Como $f(\overline{x}) > f(\overline{x}) - \varepsilon$, então $\overline{x} \in [S_{\alpha_0}(f)]^c$, onde

$\alpha_0 = f(\overline{x}) - \varepsilon$. Assim, existe $r > 0$ tal que $B(\overline{x}, r) \subset [S_{\alpha_0}(f)]^c$, isto é, $f(x) > f(\overline{x}) - \varepsilon, \forall x \in B(\overline{x}, r)$ ou, equivalentemente, f é sci em $\overline{x}$. □

A seguinte proposição nos fornece condições suficientes para que o conjunto factível de (P) seja fechado.

Corolário 2.6 *Seja $U \subset \mathbb{R}^n$ um conjunto fechado. Se as restrições $g_i, i \in I$, são sci, então $X = \{x \in U : g_i(x) \leq 0, i \in I\}$ é fechado.*

Demonstração. Se g_i são sci, então pela Proposição 2.4, o conjunto $S_0(g_i)$ são fechados. Mas, podemos escrever

$$X = U \cap (\bigcap_{i=1}^{m} S_0(g_i)),$$

ou seja, X é interseção de $m+1$ fechados e, portanto, X é fechado. □

Um conceito necessário nesta discussão é o de função coerciva:

Definição 2.7 *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função dada e $X \subset \mathbb{R}^n$ um subconjunto não vazio. Diremos que f é coerciva sobre X se para toda sequência $(x^k) \subset X$, tal que $||x^k|| \to \infty$, tem-se $f(x^k) \to +\infty$.*

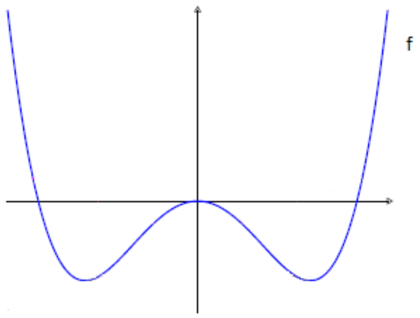


Figura 2.4: Função coerciva [8]

A Figura 2.4 ilustra que $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^4 - x^2$ é coerciva.

Proposição 2.8 *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ e $X \subset \mathbb{R}^n$ um subconjunto não vazio. Então f é coerciva sobre X se, e somente se, o conjunto*

$$S_\alpha(f, X) = \{x \in X : f(x) \leq \alpha\}$$

é limitado, para todo $\alpha \in \mathbb{R}$.

Demonstração. Inicialmente, suponha que f é coerciva e, por contradição, suponha que exista um $\alpha' \in \mathbb{R}$ tal que o conjunto $S_{\alpha'}(f, X)$ é ilimitado. Então, existe uma sequência $(x^k) \subset X$ tal que $||x^k|| \to \infty$ e $f(x^k) \leq \alpha'$, para cada k, que contradiz o fato de f ser coerciva.

Agora, demonstramos a recíproca. Tome $(x^k) \subset X$ uma sequência tal que $||x^k|| \to \infty$, e suponha que os conjuntos $S_\alpha(f, X)$ são limitados, $\forall \alpha \in \mathbb{R}$. Então, para cada $n \in \mathbb{N}$ fixado, existe um índice $k(n)$ tal que $x^k \notin S_{k(n)}(f, X), \forall k \geq k(n)$. Como $x^k \notin S_{k(n)}(f, X)$ então $f(x^k) > k(n), \forall k \geq k(n)$ e, portanto, $f(x^k) \to +\infty$. Logo, f é coerciva. □

Corolário 2.9 *Suponha que no problema (P), pelo menos uma das restrições $g_i, i \in I$ é coerciva. Então, o conjunto factível X é limitado.*

Demonstração. Suponha que existe um índice $i_0 \in I$ tal que g_{i_0} é coerciva. Então, pela Proposição anterior, o conjunto de nível $S_0(g_{i_0}) = \{x \in \mathbb{R}^n : g_{i_0}(x) \leq 0\}$ é limitado. Com maior razão, $X \subset S_0(g_{i_0})$ também é limitado. □

Outro conceito que necessitaremos é o de solução ótima com tolerância:

Definição 2.10 *Sejam $\varepsilon \geq 0$ dado e $v(P)$ o valor ótimo do problema (P). Definimos o conjunto*

$$\widehat{X}(\varepsilon) = \{x \in X : f(x) \leq v(P) + \varepsilon\}$$

ao qual chamaremos ***conjunto das soluções ótimas de (P) com tolerância*** ε.

Observe que, em particular, o conjunto $\widehat{X}(0) = \{x \in X : f(x) \leq v(P)\}$.

Exemplo 2.11 *No caso da figura abaixo,* $\hat{X}(\varepsilon) = \left\{x \in \mathbb{R}^+ : x^2 \leq \varepsilon\right\}$, *cuja função é* $f(x) = x^2$.

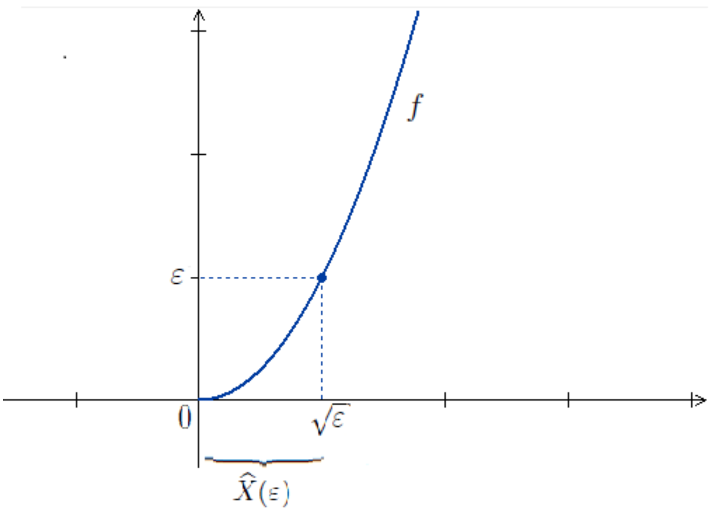


Figura 2.5: Valor ótimo com tolerância ε

O próximo teorema apresenta um resultado de existência de soluções, relaxando-se a hipótese de compacidade do conjunto factível. Para demonstrar tal fato, necessitaremos do seguinte resultado:

Proposição 2.12 *Se f é sci e X é um conjunto fechado, então os conjuntos $\widehat{X}(\varepsilon)$ são fechados, para cada $\varepsilon \geq 0$.*

Demonstração. Tome $\varepsilon \geq 0$. Se $v(P) = -\infty$, então $\widehat{X}(\varepsilon) = \varnothing$, logo é limitado. Agora, suponhamos que $v(P)$ é finito. Como f é sci, pela Proposição (2.5), o conjunto $\{x \in \mathbb{R}^n : f(x) \leq v(P) + \varepsilon\}$ é fechado. Além disso, por hipótese, X também é fechado. Portanto, o conjunto

$$\widehat{X}(\varepsilon) = X \cap \{x \in \mathbb{R}^n : f(x) \leq v(P) + \varepsilon\}$$

é fechado. □

Finalmente, estabeleceremos os principais resultados desta Seção.

Teorema 2.13 *No problema (P), suponha que o conjunto X é compacto e não vazio e que a função objetivo f é sci. Então o conjunto $\widehat{X}(0)$ das soluções ótimas de (P) é compacto e não vazio.*

Demonstração. A demonstração se dará em duas partes. Na primeira delas, verificaremos que $\widehat{X}(0)$ é não-vazio e, na segunda, que é compacto.

Seja $\beta = v(P)$ o valor ótimo do problema (P), então existe uma sequência $(x^k) \subset X$ tal que $\beta = \lim\limits_{k\to\infty} f(x^k)$ e, se $\alpha > \beta$, existe $\overline{k} \in N$ tal que

$$x^k \in \{x \in X : f(x) \leq \alpha\} = S_\alpha(f, X), \forall k \geq \overline{k}.$$

Sendo X limitado, o conjunto $S_\alpha(f, X)$ também é limitado. Assim sendo, a sequência (x^k) é limitada e portanto, admite uma subsequência $(x^{k'})$ tal que $x^{k'} \stackrel{k'\to\infty}{\longrightarrow} \overline{x}$. Como $x^{k'} \in S_\alpha(f, X), \forall k' \geq \overline{k}$ e $S_\alpha(f, X)$ é fechado, então $\overline{x} \in S_\alpha(f, X)$. Considere a sequência $(f(x^{k'}))$. É claro que $f(x^{k'}) \stackrel{k'\to\infty}{\longrightarrow} \beta$, e sendo f uma função sci,

$$f(\overline{x}) \leq \liminf_{k'\to\infty} f(x^{k'}) = \lim_{k'\to\infty} f(x^{k'}) = \beta.$$

Por outro lado, como $\overline{x} \in X$, temos que $\beta \leq f(\overline{x})$ e portanto, $f(\overline{x}) = v(P)$, ou seja, $\overline{x} \in \widehat{X}(0)$.

Agora, verifiquemos que o conjunto $\widehat{X}(0)$ é compacto. Sendo f sci e X fechado, então, pela Proposição 2.12, o conjunto $\widehat{X}(0)$ é fechado. Além disto, $\widehat{X}(0) \subset X$ e, por hipótese, X é limitado. Logo, $\widehat{X}(0)$ é limitado. Portanto $\widehat{X}(0)$ é compacto. □

Observação 1 *O Teorema 2.13, fornece um resultado de existência de soluções para o problema (P) com hipóteses relaxadas de continuidade na função objetivo f. (Compare com a Proposição 2.1).*

O próximo resultado estabelece que, sob hipóteses de coercividade da função objetivo, os conjuntos $\widehat{X}(\varepsilon)$ são limitados.

Proposição 2.14 *Seja f uma função coerciva sobre o conjunto factível X. Então, para cada $\varepsilon \geq 0$, os conjuntos $\widehat{X}(\varepsilon)$ são limitados.*

Demonstração. Tome $\varepsilon \geq 0$. Se $v(P) = -\infty$, então $\widehat{X}(\varepsilon) = \emptyset$ e, portanto, limitado. Suponhamos, que $v(P)$ é finito. Se f é coerciva sobre X,, então pela Proposição (2.8) o conjunto $S_{v(P)+\varepsilon}(f, X) = \widehat{X}(\varepsilon)$ é limitado. □

O teorema que se segue apresenta outro resultado de existência de soluções, relaxando-se a hipótese de compacidade do conjunto factível.

Teorema 2.15 *Suponha que X é fechado, não vazio e que a função f é sci. Se f é coerciva, então $\widehat{X}(0)$ é compacto e não-vazio.*

Demonstração. Seja $\beta = v(P)$ o valor ótimo de (P). Então existe uma sequência $(x^k) \subset X$ tal que $\beta = \lim\limits_{k\to\infty} f(x^k)$. Primeiro vamos demonstrar que $\widehat{X}(0)$ é não-vazio. Seja $\alpha > \beta$, então existe $\overline{k} \in N$ tal que

$$x^k \in S_\alpha(f, X) = \{x \in X : f(x) \leq \alpha\}, \forall k \geq \overline{k}.$$

Segue da Proposição 2.8 que o conjunto $S_\alpha(f,X)$ é limitado e, existe uma subsequência $(x^{k'})$ de (x^k) tal que $x^{k'} \overset{k'\to\infty}{\longrightarrow} \overline{x}$. Como X é fechado, temos $\overline{x} \in X$. Considere a sequência $(f(x^{k'}))$. Temos que $\beta = \lim\limits_{k\to\infty} f(x^{k'})$ e como f é sci,

$$f(\overline{x}) \leq \liminf_{k'\to\infty} f(x^{k'}) = \beta.$$

Como $\overline{x} \in X$, temos que $\beta \leq f(\overline{x})$, ou seja $\beta = f(\overline{x})$. Logo, $\overline{x} \in \widehat{X}(0)$.

Provaremos agora a compacidade de $\widehat{X}(0)$. Como X fechado e f sci, segue da Proposição (2.12) que $\widehat{X}(0)$ é fechado. Sendo f coerciva, então pela Proposição 2.14 $\widehat{X}(0)$ é limitado.

□

Exemplo 2.16 *[13] Dados uma região convexa fechada não-vazia $C \subseteq \mathbb{R}^n$ e um ponto $\overline{x} \in \mathbb{R}^n$, aplicaremos os teoremas vistos anteriormente sobre o estudo da existência e caracterização dos pontos em C que minimizam, sobre C, uma certa distância euclideana ao ponto $\overline{x}$, isto é, o estudo do problema:*

$$\begin{array}{l} \textit{Minimizar } \ \|x-\overline{x}\|_Q \\ \textit{sujeito a: } \ \ x \in C \end{array} \tag{Q}$$

onde $\|y\|_Q = \left(y^TQy\right)^{\frac{1}{2}}$, Q é uma matriz $(n \times n)$ simétrica definida positiva (em particular, se Q=I, I a matriz identidade de ordem n, então temos a distância euclideana habitual).

Mais precisamente, estudaremos o problema equivalente, com função objetivo diferenciável:

$$\begin{array}{l} \textit{Minimizar } \frac{1}{2}\|x-\overline{x}\|_Q^2 \\ \textit{sujeito a: } \ \ x \in C \end{array} \tag{Q}$$

A função objetivo $f(x)$ de (Q) é contínua. Mostraremos que $f(x)$ é coerciva: Como Q é definida positiva, então a seguinte relação é válida para todo $x \in \mathbb{R}^n$

$$x^TQx \geq \alpha\|x\|^2, \ \textit{com } \alpha > 0$$

Logo,

$$(\alpha)^{\frac{1}{2}}\|x\| \leq \|x\|_Q \leq \|x-\overline{x}\|_Q + \|\overline{x}\|_Q \tag{2.4}$$

onde $\alpha > 0$ e $\|x\| \to +\infty$, por hipótese. Então, por (2.4), $\|x-\overline{x}\|_Q \to +\infty$. Como $f(x) = \frac{1}{2}\|x-\overline{x}\|_Q^2$, temos que $f(x) \to +\infty$. Portanto $f(x)$ é coerciva e, pelo Teorema 2.15, concluímos que o problema (Q) admite soluções ótimas.

Capítulo 3

Condições necessárias de otimalidade

Neste Capítulo, consideraremos as condições que devem ser satisfeitas para que um ponto factível do problema (P) seja ótimo. Serão consideradas condições de 1^a. ordem (i.e., condições que envolvem as derivadas de 1^a. ordem) e também de segunda ordem (i.e., condições que envolvem as derivadas segundas).

3.1 Condições de primeira ordem

Iniciamos com uma caracterização geométrica de otimalidade local para o problema (P):

Proposição 3.1 *Suponha que as funções f, g_i, $i \in I$, são diferenciáveis. Se $x^* \in X$ é um minimizador local de (P), então o sistema*

$$\begin{array}{l} \nabla^T f(x^*)d < 0 \\ \nabla^T g_i(x^*)d < 0, i \in I(x^*) \end{array} \tag{3.1}$$

não admite solução $d \in \mathbb{R}^n$.

Demonstração. Seja $x^* \in X$ um minimizador local de (P). Suponhamos por contradição que exista uma direção $\widehat{d} \in \mathbb{R}^n$ que resolve o Sistema (3.1). Como $x^* \in U$ e U é aberto, então existe um $\delta_0 > 0$ tal que $x^* + t\widehat{d} \in U$, para todo $t \in (0, \delta_0)$. Além disso, para cada $i \in I(x^*)$ fixado, temos

$$0 > \nabla^T g_i(x^*)\widehat{d} = \lim_{t \to 0^+} \frac{g_i(x^* + t\widehat{d}) - g_i(x^*)}{t}.$$

Logo, para cada $i \in I(x^*)$ existe $\delta_i > 0$ tal que

$$g_i(x^* + t\widehat{d}) - g_i(x^*) = g_i(x^* + t\widehat{d}) < 0, \forall t \in (0, \delta_i).$$

Por outro lado, para $i \in I \setminus I(x^*)$ temos que $g_i(x^*) < 0$ e como as funções g_i são contínuas, existe $\delta_i > 0$ tal que

$$g_i(x^* + t\widehat{d}) < 0, \forall t \in (0, \delta_i).$$

Definindo-se $\widehat{\delta} = \min\{\delta_0, \cdots, \delta_m\}$, temos que $x^* + t\widehat{d} \in X$, para todo $t \in (0, \widehat{\delta})$. Por outro lado, temos

$$0 > \nabla^T f(x^*)\widehat{d} = \lim_{t \to 0^+} \frac{f(x^* + t\widehat{d}) - f(x^*)}{t}$$

e portanto, existe $\delta' > 0$ tal que $f(x^* + t\widehat{d}) < f(x^*), \forall t \in (0, \delta')$.

Defina $\delta = \min\{\widehat{\delta}, \delta'\}$. Para cada $t \in (0, \delta)$, obtemos os pontos da forma $x^* + t\widehat{d}$ que são factíveis para (P) e $f(x^* + t\widehat{d}) < f(x^*)$, contrariando a otimalidade local de x^*. $\square$

Utilizando-se um Teorema de Alternativa, podemos converter a condição geométrica de otimalidade dada na proposição anterior em uma regra de multiplicadores.

Adotaremos a seguinte notação: Para $x, y \in \mathbb{R}^n$, escreveremos

$$\begin{aligned} x &< y \iff x_i < y_i,\; i = 1, \cdots, n \\ x &\leqq y \iff x_i \leq y_i,,\; i = 1, \cdots, n \\ x &\leq y \iff x \leqq y \text{ e } x \neq y. \end{aligned}$$

Observação 2 *Recordemos que:*

(i) Um conjunto $D \subset \mathbb{R}^n$ é chamado ***conjunto convexo*** *se para quaisquer $x \in D, y \in D$ e $\alpha \in [0, 1]$ tem-se que $\alpha x + (1 - \alpha)y \in D$.*

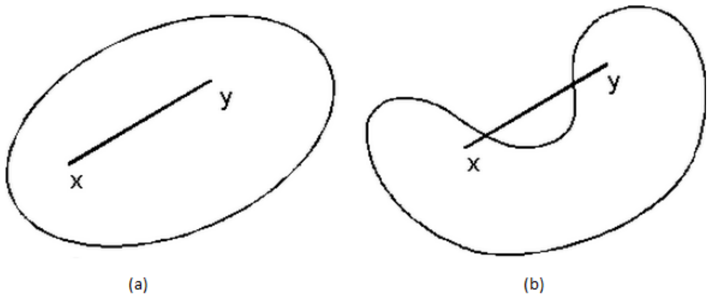


Figura 3.1: (a) Conjunto convexo (b) Conjunto não convexo

(ii) Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Dizemos que a função $f : D \to \mathbb{R}$ é convexa em D quando para quaisquer $x, y \in D$ e $t \in [0, 1]$, tem-se que

$$f((1 - t)x + ty) \leqq \; (1 - t)f(x) + tf(y)$$

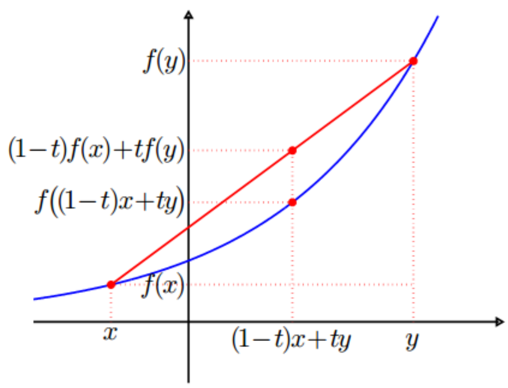


Figura 3.2: Função convexa

Pode-se provar: se $f : D \to \mathbb{R}$ é uma função convexa e continuamente diferenciável em D, convexo, então $\forall x, y \in D$

$$f(x) + \nabla^T f(x)(y - x) \leqq f(y).$$

Para maiores detalhes, veja [1].

Lema 3.2 *[Teorema de Gordan Generalizado] Sejam $D \subseteq \mathbb{R}^n$ um conjunto convexo e $f : D \to \mathbb{R}^p$, uma função convexa definida em D. Temos que:*

(i) *Não existe $x \in D$ tal que:*

$$f_j(x) < 0, \ \forall j \in J$$

se, e somente se

(ii) *Existem m escalares não-negativos μ_j, $j \in J$, não todos nulos, tais que:*

$$\sum_{j=1}^{p} \mu_j f_j(x) \geqq 0, \ \forall x \in D$$

em que $J = \{1, \ldots, p\}$.

Demonstração. $(ii) \Rightarrow (i)$. Suponha por contradição que (ii) ocorre, entretanto existe $\overline{x} \in D$ tal que:

$$f_j(\overline{x}) < 0, \ \forall j \in J \tag{3.2}$$

Como (ii) é válido por hipótese, temos que:

$$\begin{aligned} &\sum_{j=1}^{p} \mu_j f_j(x) \geqq 0, \ \forall x \in D \\ &\mu_j \geqq 0 \\ &\mu_{j_0} > 0, \textit{para algum } j_0 \in J. \end{aligned} \tag{3.3}$$

Assim, de (3.2) e (3.3) concluímos que

$$\sum_{j=1, j \neq j_0}^{p} \mu_j f_j(\overline{x}) + \mu_{j_0} f_{j_0}(\overline{x}) < 0$$

o que é um absurdo. Portanto, $(ii) \Rightarrow (i)$.

Na sequência abaixo, mostraremos que $(i) \Rightarrow (ii)$.

- Seja $\phi = \{y \in \mathbb{R}^m : f(x) \leqq y, (\textit{i. é.}, f_j(x) \leqq y_j, \ j \in J), \ \textit{para algum } x \in D\}$, de modo que o conjunto ϕ seja não vazio e convexo.

- Seja o conjunto convexo $K = \{y \in \mathbb{R}^m : y_j < 0, j \in J\}$. Ou seja, o ortante negativo do $\mathbb{R}^m$.

- Mostremos que $K \cap \phi = \emptyset$. Suponha que exista $\overline{y} \in K \cap \phi$. Então, existe $\overline{x} \in D$ tal que $f_j(\overline{x}) \leq \overline{y} < 0, \ \forall j \in J$, o que impossível por (i).

- Agora, dado que K é convexo, ϕ é convexo e $K \cap \phi = \emptyset$, podemos aplicar o *Teorema da Separação*, (para maiores detalhes, veja [22]), o qual garante a existência de $h \in \mathbb{R}^m$ $(h \neq 0)$ e $\alpha \in \mathbb{R}$, tais que:

$$h \cdot y \geqq \alpha \geqq h \cdot z, \ \forall y \in \phi \ e \ \forall z \in K, \text{onde } h.y \text{ é o produto interno usual}$$

- Mostremos que $h \geqq 0$. Suponha por contradição que existe $j_0 \in J$ tal que $h_{j_0} < 0$. Defina $z(\rho) \in K$ (com $\rho \in \mathbb{R}_-$), como:

$$z_j(\rho) = \begin{cases} \frac{1}{\rho}, & se\ j \neq j_0 \\ \rho, & se\ j = j_0 \end{cases}$$

então, de

$$\begin{aligned} & h \cdot z \leqq \alpha\ (\alpha \in \mathbb{R}), \forall z \in K \\ & z(\rho) \in K,\ \forall \rho \in \mathbb{R}_- \\ & h_{j_0} < 0 \end{aligned}$$

concluímos que $0 < \rho \cdot h_{j_0} \leqq \alpha,\ \forall \rho < 0$ onde $\alpha \in \mathbb{R}$. O que é impossível. Logo, $h \geqq 0$.

- Por um raciocínio similar, temos que $\alpha \geqq 0$. Assim, definindo $\mu_j = h_j,\ j \in J$, temos o resultado.

□

Em particular, temos o conhecido Teorema de Gordan:

Corolário 3.3 *[Teorema de Gordan] Se $\{a_j\}_{j=1}^m$ é um conjunto dado de p vetores no $\mathbb{R}^n$, então:*

(i) *Não existe $x \in \mathbb{R}^n$ tal que:*

$$a_j \cdot x < 0,\ \forall j \in J \tag{3.4}$$

se, e somente se,

(ii) *Existe m escalares não negativos, $\mu_j,\ j \in J$, não todos nulos, tais que:*

$$\sum_{j=1}^{m} \mu_j \cdot a_j = 0$$

Aplicando o Teorema de Gordan ao sistema de desigualdade (3.1) obtemos o Teorema de Fritz-John.

Teorema 3.4 *[Fritz-John] Seja $x^* \in X$ um minimizador local de (P). Suponha que as funções f, g_i $(i \in I(x^*))$ são diferenciáveis e que as funções g_i $(i \in I \backslash I(x^*))$ são contínuas. Então, existem escalares $\lambda_0, \lambda_i \in \mathbb{R}$, $i \in I$, não todos nulos, tais que,*

$$\lambda_0 \nabla f(x^*) + \sum_{i \in I} \lambda_i \nabla g_i(x^*) = 0 \tag{3.5}$$

$$\lambda_0, \lambda_i \geqq 0,\ i \in I \tag{3.6}$$

$$\lambda_i g_i(x^*) = 0,\ i \in I. \tag{3.7}$$

Demonstração. Seja x^* uma solução local de (P). Pela Proposição 3.1, o Sistema (3.1) não possui solução $d \in \mathbb{R}^n$. Então, aplicando o Teorema de Gordan (Corolário 3.3) ao Sistema (3.1), existem $\widehat{\lambda}_i \geqq 0$, $i \in \{0\} \cup I(x^*)$, não todos nulos e tais que

$$\widehat{\lambda}_0 \nabla^T f(x^*)d + \sum_{i \in I(x^*)} \widehat{\lambda}_i \nabla^T g_i(x^*)d = 0, \forall d \in \mathbb{R}^n$$

isto é,

$$\widehat{\lambda}_0 \nabla f(x^*) + \sum_{i \in I(x^*)} \widehat{\lambda}_i \nabla g_i(x^*) = 0.$$

Assim, basta definir:

$$\lambda_i = \begin{cases} \widehat{\lambda}_i, i \in I(x^*) \cup \{0\} \\ 0, \ i \in I \setminus I(x^*). \end{cases}$$

□

Podemos observar que no Teorema 3.4, não existe a garantia de que o multiplicador associado à função objetivo λ_0 seja não nulo. Dessa forma, se este multiplicador for zero, perdemos as informações contidas na derivada da função objetivo. Para contornar este inconveniente, inserimos nas hipóteses das condições necessárias, alguma condição de regularidade (ou qualificação de restrição).

Por simplicidade, admitiremos a seguinte qualificação de restrição:

Definição 3.5 *Diremos que o problema (P) satisfaz a qualificação de restrição* ***(QR)*** *em x^* se o conjunto $\{\nabla g_i(x^*)\}_{i \in I(x^*)}$ for linearmente independente.*

Se às hipóteses do Teorema de Fritz-John acrescentar a qualificação de restrição (QR), o multiplicador $\lambda_0 \neq 0$.

Teorema 3.6 *[Kuhn-Tucker] Suponha que $x^* \in X$ é um minimizador local de (P), que as funções f, g_i $(i \in I(x^*))$ são diferenciáveis e que as funções g_i $(i \in I \setminus I(x^*))$ são contínuas. Além disto, suponha que as restrições do problema (P) satisfazem a qualificação de restrição (QR) no ponto x^*. Então, existem $\mu_i \in \mathbb{R}$, $i \in I$ tais que*

$$\nabla f(x^*) + \sum_{i \in I} \mu_i \nabla g_i(x^*) = 0 \tag{3.8}$$

$$\mu_i \geqq 0, \ i \in I \tag{3.9}$$

$$\mu_i g_i(x^*) = 0, \ i \in I. \tag{3.10}$$

Demonstração. Pelo Teorema (3.4), existem multiplicadores λ_0, λ_i $(i \in I)$ tais que se verificam as equações (3.5)-(3.7). Se $\lambda_0 = 0$, por (3.5) teríamos $\sum_{i \in I} \lambda_i \nabla g_i(x^*) = 0$, com $\lambda_i \geqq 0$ e não todos nulos, o que contraria a condição (QR). Basta definir

$$\mu_i = \frac{\lambda_i}{\lambda_0}, i \in I$$

e obtemos o resultado desejado. □

Observação 3 *Observe que as condições de Kuhn-Tucker (Teorema 3.6) podem ser obtidas com qualificações de restrição mais gerais que a condição (QR). Para maiores detalhes, veja, por exemplo [24, 9, 5, 17, 18, 32, 23, 31].*

Definição 3.7 *Um ponto $x^* \in X$ factível para (P) é dito ser um* ***ponto Kuhn-Tucker*** *se existem $\mu_i \in \mathbb{R}$, $i \in I$, satisfazendo às Equações (3.8)-(3.10).*

Exemplo 3.8 *Seja o seguinte problema:*

$$\begin{array}{ll} \textit{Minimizar } f(x) & \\ \textit{sujeito a: } & \begin{array}{l} g_i(x) \leqq 0, i = 1, \cdots, 4 \\ x \in U \subseteq \mathbb{R}^2 \end{array} \end{array} \qquad \text{(M)}$$

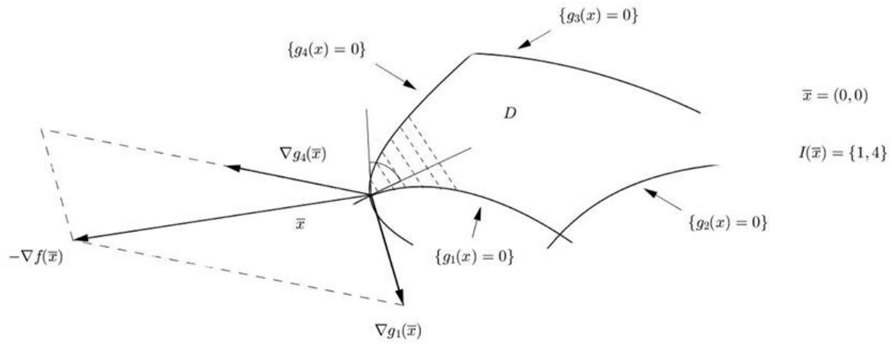


Figura 3.3: Condições de Kuhn-Tucker

Na Figura (3.3), traçamos as quatro restrições nos casos em que $g_i(x) = 0, i = 1, \cdots, 4$ com $x \in \mathbb{R}^2$. O conjunto das restrições ativas em $\overline{x}$ é dado por $I(\overline{x}) = \{1, 4\}$.

Os gradientes $\nabla g_i(x)$ são naturalmente ortogonais as respectivas funções $g_i(x)$ no ponto $\overline{x}$, e apontam para a direção de máximo crescimento de $g_i(x)$ a partir de $\overline{x}$ (onde $g_i(x) = 0$).

Se o ponto $\overline{x}$ é mínimo local do problema diferenciável (M), o Teorema de Kuhn-Tucker (3.6) garante a existência dos multiplicadores $\mu_1, \mu_4 \geq 0$, tais que,

$$-\nabla f(\overline{x}) = \mu_1 \nabla g_1(\overline{x}) + \mu_4 \nabla g_4(\overline{x})$$

isto é, $-\nabla f(\overline{x})$ pode ser escrito como combinação linear (com coeficientes positivos) dos gradientes das restrições ativas em $\overline{x}$.

3.2 Condições de segunda ordem

Nesta seção, obteremos as condições necessárias de otimalidade para o problema (P) utilizando as informações das derivadas segundas das funções do problema. Ao longo desta seção, suporemos que as funções f, g_i são duas vezes diferenciável.

Para isto, necessitaremos do conceito de direção crítica abaixo.

Definição 3.9 *Dizemos que o vetor $d \in \mathbb{R}^n$ é uma direção crítica do problema (P) no ponto $x^* \in X$ se*

$$\begin{array}{l} \nabla^T f(x^*)d \leq 0 \\ \nabla^T g_i(x^*)d \leq 0, i \in I(x^*) \end{array}$$

e denotaremos por $D(x^*)$ o conjunto de direções críticas em x^*.

E do conhecido Teorema da Alternativa abaixo.

Lema 3.10 *[Teorema de Motzkin] Sejam A, C e D matrizes dadas, com $A \neq 0$. Então, exatamente uma das seguintes condições deve ocorrer:*

(I) *O sistema*

$$\left.\begin{array}{l} Ax > 0 \\ Cx \geqq 0 \\ Dx = 0 \end{array}\right\}$$

tem uma solução x;

(II) *O sistema*

$$\left.\begin{array}{l} A^T y_1 + C^T y_2 + D^T y_3 = 0 \\ y_1 \geq 0, y_2 \geqq 0 \end{array}\right\}$$

tem solução (y_1, y_2, y_3).

Para maiores detalhes e prova, veja [22].

Teorema 3.11 *Seja $x^* \in X$ um minimizador local de (P). Suponha que as funções f, g_i $(i \in I(x^*))$ são duas vezes continuamente diferenciáveis e que as funções g_i $(i \in I \setminus I(x^*))$ são contínuas. Então, para cada direção crítica $d \in D(x^*)$, existem multiplicadores $\lambda_0, \lambda_i \geq 0$ $(i \in I)$ tais que*

$$\lambda_0 \nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0 \tag{3.11}$$

$$\lambda_i g_i(x^*) = 0, i \in I \tag{3.12}$$

$$\lambda_0 \nabla^T f(x^*) d = 0 \tag{3.13}$$

$$\lambda_i \nabla^T g_i(x^*) d = 0, i \in I(x^*) \tag{3.14}$$

$$d^T (\nabla^2 f(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla^2 g_i(x^*)) d \geqq 0. \tag{3.15}$$

Além disto, se (QR) vale em x^*, então para todo $d \in D(x^*)$, teremos $\lambda_0 \neq 0$.

Para facilitar a demonstração, defina o seguinte conjunto:

$$I_0 = \left\{ i \in I(x^*) : \nabla^T g_i(x^*) d = 0 \right\}$$

Observe que este conjunto de índices é equivalente as afirmações (3.12) e (3.14).
Demonstração. Seja d uma direção crítica fixa e arbitrária. Então, o sistema

$$\left\{ \begin{array}{l} \nabla^T f(x^*) z + d^T (\nabla^2 f(x^*)) d < 0 \\ \nabla^T g_i(x^*) z + d^T (\nabla^2 g_i(x^*)) d \leqq 0, \; i \in I_0 \end{array} \right. \tag{3.16}$$

não tem solução $z \in \mathbb{R}^n$. Isto é equivalente a

$$\begin{cases} \nabla^T f(x^*)z + [d^T(\nabla^2 f(x^*))d]t < 0 \\ \nabla^T g_i(x^*)z + [d^T(\nabla^2 g_i(x^*))d]t \leqq 0, \ i \in I_0 \\ -t < 0, \end{cases}$$

o qual não tem solução $z \in \mathbb{R}^n$, $t \in \mathbb{R}$. Pelo Teorema de Motzkin (Lema 3.10), existem multiplicadores $\xi, \lambda_0 \in \mathbb{R}$ e $\lambda_i \in \mathbb{R}^m$ tais que

$$\begin{cases} \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0 \\ d^T(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*))d - \xi = 0 \\ (\lambda_0, \xi) \geq 0, \lambda_i \geqq 0, \lambda_i = 0, \forall i \notin I_0 \end{cases}$$

Como $(\lambda_0, \xi) \geq 0$, temos que: $\lambda_0 \geq 0$ e $\xi \geqq 0$, ou, $\lambda_0 \geqq 0$ e $\xi > 0$. E assim, existem $\lambda_0 \in \mathbb{R}$ e $\lambda_i \in \mathbb{R}^m$ tais que

$$\begin{cases} \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0 \\ d^T(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*))d > 0 \\ \lambda_0 \geqq 0, \lambda_i \geqq 0, \lambda_i = 0, \forall i \notin I_0 \end{cases} \tag{3.17}$$

ou,

$$\begin{cases} \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0 \\ d^T(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*))d \geqq 0 \\ \lambda_0 \geq 0, \lambda_i \geqq 0, \lambda_i = 0, \forall i \notin I_0. \end{cases} \tag{3.18}$$

Assuma que (3.18) não ocorre. Isto é equivalente a

$$\begin{cases} \lambda_0 \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0 \\ d^T(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*))d - \xi = 0 \\ \lambda_0 \geq 0, \xi \geqq 0, \lambda_i \geqq 0, \lambda_i = 0, \forall i \notin I_0. \end{cases} \tag{3.19}$$

Logo, novamente pelo Teorema de Motzkin (Lema 3.10), existem $z, t \geqq 0$, satisfazendo

$$\begin{cases} d^T[\nabla f(x^*)z + \nabla^2 f(x^*)]dt < 0 \\ d^T[(\nabla g_i(x^*)z + \nabla^2 g_i(x^*)]dt \leqq 0, \ \forall i \in I_0. \end{cases}$$

Mas, como (3.16) não tem solução, temos que $t = 0$. Por isso,

$$\begin{cases} \nabla^T f(x^*)z < 0 \\ \nabla^T g_i(x^*)z \leqq 0, \ \forall i \in I_0. \end{cases}$$

Por outro lado,

$$\begin{cases} \nabla^T f(x^*)d = 0 \\ \nabla^T g_i(x^*)d = 0, \ \forall i \in I_0 \end{cases}$$

e

$$\begin{cases} \nabla^T f(x^*)d < 0 \\ \nabla^T g_i(x^*)d < 0, \ \forall i \in I \setminus I_0 \end{cases}$$

porque d é crítica. Assim, vale que

$$\begin{cases} \nabla^T f(x^*)(d + \varepsilon z) < 0 \\ \nabla^T g_i(x^*)(d + \varepsilon z) \leqq 0, \ \forall i \in I \end{cases}$$

para qualquer $\varepsilon > 0$ suficientemente pequeno. O que contradiz (3.11). E isto completa a prova. □

Definição 3.12 *Um ponto $x^* \in X$ factível para (P) que satisfaz as condições (3.11)-(3.15), com $\lambda_0 \neq 0$, para toda direção crítica $d \in D(x^*)$ é chamado* ***ponto Kuhn-Tucker de segunda ordem****.*

No próximo capítulo estabeleceremos condições adicionais para que pontos Kuhn-Tucker e Kuhn-Tucker de segunda ordem sejam minimizadores locais.

Capítulo 4

Condições Suficientes de Otimalidade

Neste capítulo abordaremos critérios que garantem que as condições necessárias vistas no capítulo anterior também são suficientes para a otimalidade local. Tais resultados serão obtidos para funções convexas e posteriormente generalizaremos a uma classe mais ampla de funções.

É bastante conhecido que sob hipóteses de convexidade as condições de Kuhn-Tucker são suficientes para a otimalidade global.

Teorema 4.1 *Seja $\overline{x} \in X$ um ponto factível de (P) e suponha que as funções f, g_i ($i \in I(\overline{x})$) são continuamente diferenciáveis e convexas. Se existe $\lambda \in \mathbb{R}^n$ que satisfaz as condições*

$$\nabla f(\overline{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\overline{x}) = 0 \tag{4.1}$$

$$\lambda_i \geqq 0, i \in I \tag{4.2}$$

$$\lambda_i g_i(\overline{x}) = 0, i \in I, \tag{4.3}$$

então $\overline{x}$ é uma solução ótima global de (P).

Demonstração. Dado $x \in X$, temos que

$$g_i(x) \leqq 0 = g_i(\overline{x}), \forall i \in I(\overline{x})$$

e pela convexidade das g_i, temos que

$$\nabla^T g_i(\overline{x})(x - \overline{x}) \leqq 0, i \in I(\overline{x}), \tag{4.4}$$

e de (4.4) e (4.1) segue que

$$\nabla^T f(\overline{x})(x - \overline{x}) = -\sum_{i \in I} \lambda_i \nabla^T g_i(\overline{x})(x - \overline{x}) \geqq 0. \tag{4.5}$$

Mas, sendo f convexa, temos

$$\nabla^T f(\overline{x})(y - \overline{x}) \leqq f(y) - f(\overline{x}), \forall y \in X. \tag{4.6}$$

De (4.5) e (4.6), segue que

$$f(\overline{x}) \leqq f(y), \forall y \in X,$$

ou seja, $\overline{x}$ é minimizador global de (P). □

O objetivo central deste capítulo é estabelecer a suficiência das condições de Kuhn-Tucker para uma classe mais ampla de funções.

4.1 Funções invexas e suficiência das condições de Kuhn-Tucker

A suficiência da condições de Kuhn-Tucker é verificada sob hipóteses de convexidade, como observamos no início deste capítulo. Neste momento, surge um questionamento pertinente: existem classes de funções que generalizam a classe das funções convexas e que, para tais funções, as condições de Kuhn-Tucker ainda resultam suficientes para a otimalidade?

Hanson [16] foi um dos primeiros a dar contribuições significativas no intuito de enfraquecer as hipóteses de convexidade na obtenção de condições suficientes de otimalidade. Ele introduziu o conceito de função invexa e demonstrou que - para problemas irrestritos - uma função é invexa se, e somente se, todos os seus pontos estacionários são minimizadores globais. Além disso, sob hipóteses de invexidade se verifica a suficiência das condições de Kuhn-Tucker para a otimalidade global.

Inicialmente, trataremos o problema irrestrito

$$\begin{array}{l}\text{Minimizar } f(x)\\ \text{sujeito a: } x \in U\end{array} \tag{P*}$$

onde $U \subset \mathbb{R}^n$ é um subconjunto aberto e não vazio de $\mathbb{R}^n$ e $f : \mathbb{R}^n \to \mathbb{R}$ é uma função diferenciável definida em U.

O conceito de função invexa é a generalização natural do conceito de função convexa continuamente diferenciável.

Definição 4.2 *Dizemos que uma função f é* ***invexa em*** $\overline{x} \in U$ *se, para cada $x \in U$, existe uma função $\eta : U \times U \to \mathbb{R}^n$ tal que*

$$f(x) - f(\overline{x}) \geqq \nabla^T f(\overline{x})\eta(x, \overline{x}).$$

Se f for invexa em cada ponto de U, diremos simplesmente que f é ***invexa****.*

Observe que esta é a extensão natural do conceito de função convexa diferenciável. De fato, se f é convexa e continuamente diferenciável, então f é invexa, com $\eta(x, \overline{x}) = x - \overline{x}$.

Note que é possível "construir" funções invexas como a composta $F \circ \phi$, onde $F : \mathbb{R} \to \mathbb{R}$ é uma função convexa e $\phi : \mathbb{R}^n \to \mathbb{R}$ é uma função diferenciável, com derivada inversível. De fato, se $f = F \circ \phi$ e $x, y \in \mathbb{R}^n$, temos:

$$\begin{aligned} f(y) - f(x) &= F(\phi(y)) - F(\phi(x)) \\ &\geqq F'(\phi(x))[\phi(y) - \phi(x)] \\ &= f'(x)[\phi'(x)]^{-1}[\phi(y) - \phi(x)] \end{aligned}$$

Desta desigualdade, segue que f é invexa, com $\eta(y, x) = [\phi'(x)]^{-1}[\phi(y) - \phi(x)]$.

Proposição 4.3 *Seja $f : U \subset \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável. Então f é invexa se, e somente se, seus pontos estacionários são minimizadores globais de (P^*).*

Observação 4 *Lembremos que $\overline{x} \in U$ é chamado ponto estacionário (ou crítico) do Problema (P^*) se, $\nabla f(\overline{x}) = 0$.*

Demonstração. Suponha que $\overline{x} \in U$ é um ponto estacionário de f. Então $\nabla f(\overline{x}) = 0$. Da definição de invexidade, é imediato que $f(x) \geq f(\overline{x})$, para cada $x \in U$.

Agora, suponha que todo ponto estacionário de f é um minimizador global de (P*). Neste caso, basta definir:

$$\eta(x, \overline{x}) = \begin{cases} \dfrac{f(x) - f(\overline{x})}{||\nabla f(\overline{x})||^2} \nabla f(\overline{x}), \text{ se } \nabla f(\overline{x}) \neq 0 \\ 0, \text{ se } \nabla f(\overline{x}) = 0. \end{cases}$$

Por construção, $f(x) - f(\overline{x}) \geqq \nabla^T f(\overline{x}) \eta(x, \overline{x})$, $\forall x, \overline{x} \in U$, isto é, f é invexa. □

Corolário 4.4 *Se f não tem pontos estacionários, então f é invexa.*

Observação 5 *Considere a função $f(x) = 1 - \exp(-x^2)$. f não é convexa (veja figura abaixo). Mas, neste caso $f'(x) = 2x \exp(-x^2)$. Assim, seu único ponto estacionário é $x = 0$. E este ponto é minimizador global. Concluímos pela Proposição 4.3 que f é invexa.*

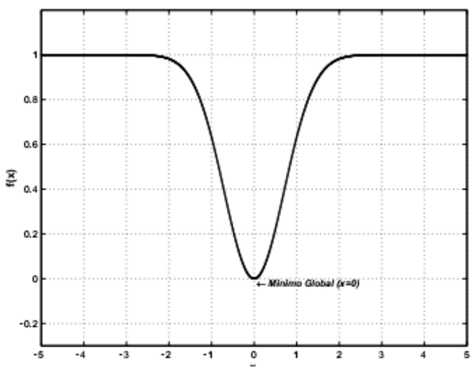


Figura 4.1: Função invexa

Observação 6 *O Corolário 4.4 é interessante, pois estabelece um critério para determinarmos se uma função é invexa sem explicitarmos a função η.*

Observação 7 *Se para o problema (P) tivermos que $\overline{x} \in X$ e as funções f, g_i, $i \in I(\overline{x})$ são todas invexas com respeito a uma mesma η, então as condições de Kuhn-Tucker (4.1)-(4.3) são suficientes para que $\overline{x} \in X$ seja um minimizador global de (P). (A prova é análoga à prova do Teorema 4.1, trocando-se o termo $x - \overline{x}$ por $\eta(x, \overline{x})$.)*

4.2 KT-invexidade e condições suficientes de primeira ordem

O objetivo desta seção é definir o chamado problema KT-invexo. Esta classe de problemas é a mais ampla classe de problemas para os quais as condições de Kuhn-Tucker são necessárias e suficientes para a otimalidade global.

Antes de definirmos KT-invexidade, vamos demonstrar que se o problema (P) for invexo (isto é, se todas as funções do problema forem invexas com respeito a uma mesma η) vale a suficiência das condições de Kuhn-Tucker para a otimalidade global.

Proposição 4.5 *Seja $\overline{x} \in X$ é um ponto Kuhn-Tucker do problema (P) e suponha que as funções f, g_i, $i \in I$ são invexas com respeito a uma mesma η. Então $\overline{x}$ é um minimizador global de (P).*

Demonstração. Seja $\overline{x} \in U$ um ponto Kuhn-Tucker de (P). Por factibilidade,

$$g(\overline{x}) \leqq 0$$

e das condições de Kuhn-Tucker, existe $\lambda \in \mathbb{R}^m$ tal que

$$\begin{aligned} \nabla f(\overline{x}) + \lambda^T \nabla g(\overline{x}) &= 0 \\ \lambda &\geqq 0 \end{aligned}$$

de modo que a seguinte condição vale: Se $g_i(\overline{x}) < 0$ então $\lambda_i = 0$.

Suponha que (P) é invexo, isto é, existe $\eta : U \times U \to \mathbb{R}^n$ tal que

$$\begin{aligned} f(x) - f(\overline{x}) &\geqq \nabla^T f(\overline{x})\eta(x, \overline{x}) \\ g_i(x) - g_i(\overline{x}) &\geqq \nabla^T g_i(\overline{x})\eta(x, \overline{x}), i \in I \end{aligned}$$

e portanto,

$$\begin{aligned} f(x) - f(\overline{x}) &\geqq \nabla^T f(\overline{x})\eta(x, \overline{x}) + 0 \\ &\geqq \nabla^T f(\overline{x})\eta(x, \overline{x}) - \lambda^T[g_i(x) - g_i(\overline{x}) + \nabla^T g_i(\overline{x})\eta(x, \overline{x})] \end{aligned}$$

e, reordenando os termos

$$\begin{aligned} f(x) - f(\overline{x}) &\geqq [\nabla f(\overline{x}) + \lambda^T \nabla g(\overline{x})]\eta(x, \overline{x}) + \lambda^T g(\overline{x}) - \lambda^T g(x) \\ &= \lambda^T g(\overline{x}) - \lambda^T g(x) = -\lambda^T g(x) \geqq 0 \end{aligned}$$

e portanto,

$$f(x) \geqq f(\overline{x})$$

para todo $x \in X$ e portanto $\overline{x}$ é minimizador global de (P). $\square$

Fazendo uma análise na demonstração da proposição acima, podemos destacar três importantes fatos, os quais nos conduzirão a um conceito de convexidade generalizada, mais fraco que a invexidade e que, ainda, preserva a suficiência das condições de Kuhn-Tucker para a otimalidade.

Primeiro, que podemos trocar $x - \overline{x}$ por $\eta(x, \overline{x})$.

Depois, olhando para a definição de invexidade é possível substituir

$$f(x) - f(\overline{x}) \geqq \nabla^T f(\overline{x})\eta(x, \overline{x}), \forall x, \overline{x} \in U$$

por

$$f(x) - f(\overline{x}) \geqq \nabla^T f(\overline{x})\eta(x, \overline{x}), \forall x, \overline{x} \in X.$$

E, por fim, é possível substituir

$$g_i(x) - g_i(\overline{x}) \geqq \nabla^T g_i(\overline{x})\eta(x, \overline{x}), i \in I$$

por

$$g_i(x) - g_i(\overline{x}) \geqq \nabla^T g_i(\overline{x})\eta(x, \overline{x}), i \in I(\overline{x})$$

o que implica

$$g_i(x) \geqq \nabla^T g_i(\overline{x})\eta(x, \overline{x}), i \in I(\overline{x}), \forall x \in X.$$

Baseando-nos nestas observações propomos a seguinte modificação (enfraquecimento) do conceito de problema invexo:

Definição 4.6 *(Martin [4]) O problema (P) é chamado* ***Kuhn-Tucker invexo*** *(ou* ***KT-invexo****, simplesmente) se existe uma função* $\eta : U \times U \to \mathbb{R}^n$ *tal que*

$$\begin{array}{rcl} f(x) - f(\overline{x}) & \geqq & \nabla^T f(\overline{x})\eta(x,\overline{x}) \\ -\nabla^T g_i(\overline{x})\eta(x,\overline{x}) & \geqq & 0, i \in I(\overline{x}) \end{array}$$

para quaisquer $x, \overline{x} \in X$.

É imediato da definição que todo problema invexo é KT-invexo. O seguinte exemplo prova que a recíproca disto é falsa, em geral.

Exemplo 4.7 *(Martin [4]) Considere o problema*

$$\begin{array}{l} \textit{Minimizar } f(x) \\ \textit{sujeito a:} \\ x \leqq 0, x \in \mathbb{R} \end{array} \qquad \text{(Q)}$$

onde suporemos que a função f *é suave, com derivada positiva. Claramente, o único ponto Kuhn-Tucker é o minimizador global* $x = 0$. *Se o Problema (Q) for invexo então* f *será convexa. De fato:*

$$\begin{array}{rcl} f(x) - f(\overline{x}) & \geqq & f'(\overline{x})\eta(x,\overline{x}) \\ -x + \overline{x} & \geqq & -\eta(x,\overline{x}), \forall x, \overline{x} \in \mathbb{R}. \end{array}$$

Como $f'(\overline{x}) > 0$, *temos que*

$$f(x) - f(\overline{x}) \geqq f'(\overline{x})(x - \overline{x}).$$

Entretanto, (Q) é KT-invexo com $\eta(x,\overline{x}) = \dfrac{f(x) - f(\overline{x})}{f'(\overline{x})}$.

Por fim, enunciamos e demonstramos e resultado fundamental desta Seção.

Teorema 4.8 *O problema (P) é KT-invexo se, e somente se, todo ponto Kuhn-Tucker é minimizador global de (P).*

Demonstração. ($\Rightarrow$) É quase imediata da definição de KT-invexidade e das observações feitas após a prova da Proposição 4.5.

($\Leftarrow$) Suponha que todo ponto Kuhn-Tucker é minimizador global de (P) e considere o par de pontos factíveis $x, \overline{x} \in X$. Se $f(x) < f(\overline{x})$, então $\overline{x}$ não é minimizador global de (P) e, por hipótese, não é um ponto Kuhn-Tucker. Isto significa que não existe $(\lambda_0, \lambda_i) \geqq 0, i \in I(\overline{x})$ com $\lambda_0 > 0$ e tal que

$$\lambda_0 \nabla f(\overline{x}) + \sum_{i \in I(\overline{x})} \lambda_i \nabla g_i(\overline{x}) = 0.$$

Logo, pelo Teorema de Alternativa de Motzkin (Lema 3.10) segue que existe um vetor $v \in \mathbb{R}^n$, que depende de $\overline{x}$ tal que

$$\begin{array}{rcl} \nabla^T f(\overline{x})v & > & 0 \\ \nabla^T g_i(\overline{x})v & > & 0, i \in I(\overline{x}). \end{array}$$

Definindo-se:

$$\eta(x,\overline{x}) = \frac{f(x) - f(\overline{x})}{\nabla^T f(\overline{x})v} v$$

temos que

$$f(x) - f(\overline{x}) = \nabla^T f(\overline{x})\eta(x,\overline{x})$$

e se $i \in I(\overline{x})$, temos $\nabla^T g_i(\overline{x})\eta(x,\overline{x}) \leqq 0$.

Agora, suponha que $f(x) \geqq f(\overline{x})$. Neste caso, definimos $\eta(x,\overline{x}) = 0$ e teremos $f(x) - f(\overline{x}) \geqq \nabla^T f(\overline{x})\eta(x,\overline{x})$ e $\nabla^T g_i(\overline{x})\eta(x,\overline{x}) = 0, i \in I(\overline{x})$. $\square$

4.3 KT-invexidade de segunda ordem e condições suficientes de segunda ordem

Nesta seção obteremos resultados semelhantes aos obtidos na seção anterior, porém considerando as condições necessárias de segunda ordem para o problema (P). Ou seja, vamos considerar uma noção de convexidade generalizada - chamada KT-invexidade de segunda ordem - a qual verifica a propriedade: (P) é KT-invexo de segunda ordem se, e somente se, todo ponto Kuhn-Tucker de segunda ordem de (P) é um minimizador global de (P). Ao longo desta seção suporemos que o problema (P) é de classe C^2.

Definição 4.9 *(Ivanov [28]) Dizemos que o problema (P) é* ***KT-invexo de segunda ordem*** *se, e somente se, existem aplicações $d : U \times U \to \mathbb{R}^n$, $\eta : U \times U \to \mathbb{R}^n$ e $\omega : U \times U \to \mathbb{R}$ tais que, para quaisquer $x, \overline{x} \in X$, $d(x,\overline{x})$ é uma direção crítica em $\overline{x}$, $\omega(x,\overline{x}) \geq 0$ e*

$$f(x) - f(\overline{x}) \geqq \nabla^T f(\overline{x})\eta(x,\overline{x}) + \omega(x,\overline{x})[d^T(x,\overline{x})\nabla^2 f(\overline{x})d(x,\overline{x})] \tag{4.7}$$

$$\nabla^T g_i(\overline{x})\eta(x,\overline{x}) + \omega(x,\overline{x})[d^T(x,\overline{x})\nabla^2 g_i(\overline{x})d(x,\overline{x})] \leqq 0, i \in I(\overline{x}). \tag{4.8}$$

Observação 8 *Todo problema KT-invexo é KT-invexo de segunda ordem. Com efeito, basta tomar $d = 0$ e $\omega = 0$ na definição de KT-invexidade de segunda ordem.*

Antes de apresentarmos um exemplo no qual mostramos que a recíproca da afirmação mencionada na Observação 7 não é verdadeira em geral, daremos uma caracterização alternativa para a KT-invexidade de 2^a ordem:

Proposição 4.10 *O problema (P) é KT-invexo de segunda ordem se, e somente se, para quaisquer $x, \overline{x} \in X$, existe uma direção crítica $d(x,\overline{x})$, um número não negativo $\omega(x,\overline{x}) \geqq 0$ e uma direção $\eta(x,\overline{x}) \in \mathbb{R}^n$ tais que a seguinte condição ocorre: Para quaisquer $x, \overline{x} \in X$ tais que $f(x) < f(\overline{x})$, então:*

(i) $\nabla^T f(\overline{x})\eta(x,\overline{x}) + \omega(x,\overline{x})[d^T(x,\overline{x})\nabla^2 f(\overline{x})d(x,\overline{x})] = -1$

(ii) $\nabla^T g_i(\overline{x})\eta(x,\overline{x}) + \omega(x,\overline{x})[d^T(x,\overline{x})\nabla^2 g_i(\overline{x})d(x,\overline{x})] \leqq 0, i \in I(\overline{x})$.

Demonstração. ($\Rightarrow$) Suponha que (P) é KT-invexo de segunda ordem. Tome $x, \overline{x} \in X$ tais que $f(x) < f(\overline{x})$ e seja:

$$p = \nabla^T f(\overline{x})\eta(x,\overline{x}) + \omega(x,\overline{x})[d^T(x,\overline{x})\nabla^2 f(\overline{x})d(x,\overline{x})]$$

temos que $p < 0$. Defina:

$$\begin{aligned} \widetilde{d}(x,\overline{x}) &= d(x,\overline{x}) \in D(\overline{x}) \\ \widetilde{\eta}(x,\overline{x}) &= \frac{1}{-p}\eta(x,\overline{x}) \in \mathbb{R}^n \\ \widetilde{\omega}(x,\overline{x}) &= \frac{\omega(x,\overline{x})}{-p} \geqq 0, \end{aligned}$$

em que d, ω e η são como especificadas na definição 4.9.

Por construção, (i) e (ii) se verificam, escolhendo-se $d(x,\overline{x}) = \widetilde{d}(x,\overline{x})$, $\eta(x,\overline{x}) = \widetilde{\eta}(x,\overline{x})$ e $\omega(x,\overline{x}) = \widetilde{\omega}(x,\overline{x})$.

($\Leftarrow$) Reciprocamente, suponha que as condições (i) e (ii) se verifiquem. Por absurdo, suporemos que o problema (P) não é KT-invexo de segunda ordem. Neste caso, existem $x, \overline{x} \in X$ tais que para todo $u \in \mathbb{R}^n$, $y \in \mathbb{R}_+$ e $z \in D(\overline{x})$, pelo menos uma das desigualdades abaixo é violada:

$$f(x) - f(\overline{x}) \geqq \nabla^T f(\overline{x})u + y[z^T \nabla^2 f(\overline{x})z] \tag{4.9}$$

$$\nabla^T g_i(\overline{x})u + y[z^T \nabla^2 g_i(\overline{x})z] \leqq 0, \forall i \in I(\overline{x}). \tag{4.10}$$

Escolhendo: $u = 0$, $y = 0$ e $z = 0$ temos que (4.10) é satisfeita. Logo, pela hipótese, (4.9) é violada. Ou seja:

$$\begin{aligned} f(x) - f(\overline{x}) &< \nabla^T f(\overline{x})u + y[z^T \nabla^2 f(\overline{x})z] \\ &= 0 \end{aligned}$$

e, portanto, $f(x) < f(\overline{x})$. Ainda, pela hipótese, existem $\eta \in \mathbb{R}^n, \omega \in \mathbb{R}_+$ e $d \in D(\overline{x})$ tais que (i) e (ii) se cumprem. Seja $t > 0$ arbitrário. Escolhendo $u = t\eta$, $y = t\omega$ e $z = d$, temos que (4.10) ocorre e, portanto, (4.9) é violada. Ou seja:

$$f(x) - f(\overline{x}) < \nabla^T f(\overline{x})(t\eta) + t\omega[d^T \nabla^2 f(\overline{x})d] < 0, \forall t > 0$$

o que é absurdo, pois $f(x) - f(\overline{x})$ é finito e $\nabla^T f(\overline{x})\eta + \omega[d^T \nabla^2 f(\overline{x})d] < 0$. □

Uma outra caracterização de ponto Kuhn-Tucker de segunda ordem é dada na seguinte proposição:

Proposição 4.11 *$x \in X$ é um ponto Kuhn-Tucker de segunda ordem se, e somente se, para toda direção crítica $d \in D(x)$ a seguinte desigualdade é válida:*

$$\min_{\eta \in \mathbb{R}^n, \omega \in \mathbb{R}_+} \{\nabla^T f(x)\eta + \omega d^T \nabla^2 f(x)d : \nabla^T g_i(x)\eta + \omega d^T \nabla^2 g_i(\overline{x})d \leqq 0, i \in I(x)\} \geqq 0 \tag{4.11}$$

Para demonstrarmos a Proposição 4.11, necessitaremos recordar alguns resultados de dualidade de problemas de Programação Linear.

Consideremos o problema linear (primal):

$$\begin{array}{l} \text{Minimizar } c^T x \\ \text{sujeito a:} \\ Ax = b \\ x \geqq 0 \end{array} \tag{PL}$$

e, a este problema, associamos o problema (dual)

$$\begin{array}{l} \text{Maximizar } b^T w \\ \text{sujeito a:} \\ wA \leqq c. \end{array} \tag{PLD}$$

Lema 4.12 *(Teorema de Dualidade) Com respeito ao par de problemas (PL) e (PLD) podemos afirmar:*

(i) Se x é factível para (PL) e w é factível para (PLD), então $c^T x \geq b^T w$.

(ii) Exatamente uma das seguintes afirmações é verdadeira:

- *Ambos problemas possuem solução ótima, x^* e w^* que verificam*

$$c^T x^* = b^T w^*.$$

- *Se um dos problemas tem valor ótimo ilimitado, então o outro é infactível.*
- *Ambos os problemas são infactíveis.*

Para maiores detalhes, veja Bazaraa [19].

Com auxílio do Lema 4.12, podemos demonstrar a Proposição 4.11.

Demonstração. (Prova da Proposição (4.11)) ($\Rightarrow$) Seja $x \in X$ um ponto Kuhn-Tucker de 2^a ordem e suponha por contradição que existam $d \in D(x), \eta \in \mathbb{R}^n$ e $\omega \in \mathbb{R}_+$ tais que

$$\nabla^T f(x)\eta + \omega[d^T \nabla^2 f(x) d] = -1 \tag{4.12}$$

$$\nabla^T g_i(x)\eta + \omega[d^T \nabla^2 f(x) d] \leqq 0, i \in I(x). \tag{4.13}$$

Como x é um ponto Kuhn-Tucker de segunda ordem, segue que existem $\lambda_0 > 0$ e $\lambda_i \geqq 0, i \in I$, os quais satisfazem as condições necessárias de segunda ordem (3.11)-(3.15). Agora, multiplicando (4.12) por λ_0 e (4.13) por λ_i e somando as equações obtidas, obtemos

$$(\lambda_0 \nabla f(x) + \sum_{i \in I(x)} \mu_i \nabla g_i(x))^T \eta + \omega(\lambda_0 d^T \nabla^2 f(x) d + \sum_{i \in I(x)} \lambda_i d^T \nabla^2 g_i(x) d) \leqq -\lambda$$

o que implica

$$\lambda_0 d^T \nabla^2 f(x) d + \sum_{i \in I(x)} \lambda_i d^T \nabla^2 g_i(x) d \leqq -\lambda < 0$$

o que é absurdo, pois $d \in D(x)$.

($\Leftarrow$) Suponha que para cada direção $d \in D(x)$ se cumpra (4.11) e, por contradição, suponha que x não é um ponto Kuhn-Tucker de 2^a ordem. Neste caso, existe uma direção crítica $d \in D(x)$, tal que, não existem multiplicadores $\lambda_0 > 0, \lambda_i \geqq 0$, $i \in I(x)$ satisfazendo as condições necessárias (3.11)-(3.15). Assim, o problema linear

$$\begin{array}{l}
\text{Maximizar } \lambda_0 \\
\text{sujeito a:} \\
\lambda_0 \nabla f(x) + \sum_{i \in I(x)} \lambda_i \nabla g_i(x) = 0 \\
\lambda_0 d^T \nabla^2 f(x) d + \sum_{i \in I(x)} \lambda_i d^T \nabla^2 g_i(x) d \geqq 0 \\
\lambda_i \geqq 0, i \in I(x)
\end{array} \tag{P1}$$

tem valor ótimo não positivo. Se tomamos $\lambda_0, \lambda_i = 0$, $i \in I(x)$, obtemos que o valor máximo deste problema é zero. Mas, o problema dual do problema (P1) é o problema (P2)

$$\begin{array}{l}
\text{Minimizar } 0 \\
\text{sujeito a:} \\
\nabla^T f(x) u - v d^T \nabla^2 f(x) d = 1 \\
\nabla^T g_i(x) u - v d^T \nabla^2 g_i(x) d \geqq 0, i \in I(x) \\
v \geqq 0.
\end{array} \tag{P2}$$

Segue do Lema 4.12 que o problema (P2) também possui solução (u, v). Portanto, existem $\eta \in \mathbb{R}^n$ e $\omega \in \mathbb{R}_+$ tais que se cumprem as equações (4.12) e (4.13): basta tomar $\eta = -u$, $\omega = v$. Desta forma, (4.11) não se cumpre para este $d \in D(x)$, o que é absurdo. $\square$

Exemplo 4.13 *(Ivanov [28]) Neste exemplo exibimos um problema que é KT-invexo de segunda ordem, mas não é KT-invexo. Considere o problema:*

$$\begin{array}{ll} \textit{Minimizar} & f(x) = \dfrac{(x_1^2+x_2^2)^2}{4} - \dfrac{x_1^2}{2} - 2x_2^2 \\ \textit{sujeito a:} & \begin{array}{l} g(x) = x_1^2 + x_2^2 - 4 \leqq 0 \\ x = (x_1, x_2) \in \mathbb{R}^2 \end{array} \end{array} \tag{R}$$

O problema (R) é KT-invexo de 2ª ordem. Verificaremos que a Proposição 4.10 é satisfeita.

Sejam $x, y \in X = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leqq 4\}$, arbitrários, com $f(y) < f(x)$. Escolhemos as funções $d = (d_1, d_2), \eta = (\eta_1, \eta_2)$ e ω como segue:

1. *Se $x = (x_1, x_2)$ é um ponto da circunferência $x_1^2 + x_2^2 \leqq 4,\ x_1 \neq 0$, então tomamos $\omega = 0,\ \eta_1 = \dfrac{-1}{3x_1},\ \eta_2 = 0,$ e $d = (0,0)$.*

2. *Se $x = (0, \pm 2)$, então a implicação (4.10) é satisfeita, pois x é minimizador global e não existe $y \in X$ tal que $f(y) < f(x)$.*

3. *Se $x_1^2 + x_2^2 < 4,\ x_2 \neq 0$, então escolhemos $\omega = 0,\ \eta_1 = 0,\ \eta_2 = \dfrac{-1}{x_1^2 x_2 + x_2^3 - 4x_2}$ e $d = (0,0)$.*

4. *Se $x = (0,0)$ então a implicação (4.10) não é satisfeita se tomarmos $\omega = 0$. Portanto, o problema (R) não é KT-invexo. Como $\nabla f(x) = 0$ podemos escolher η e d arbitrários, mas tais que $d \neq (0,0)$ e $\omega = \dfrac{1}{d_1^2 + 4d_2^2}$.*

5. *Se $x = (\pm 1, 0)$, então $\nabla f(x) = (0,0)$ e para todo par de números poderia ser escolhido para η. Tomamos d tal que $2d_1^2 < 3d_2^2$ e $\omega = \dfrac{1}{3d_2^2 - 2d_1^2}$. Porque d é crítica, se $x = (1,0)$, então tomamos $d_1 \leqq 0$ e se $x = (-1,0)$, então tomamos $d_1 \geqq 0$.*

6. *Se $x_1^2 + x_2^2 < 4,\ x_1 \neq 0, \pm 1,\ x_2 = 0$, então escolhemos $\omega = 0,\ \eta_1 = \dfrac{-1}{x_1^3 - x_1},\ \eta_2 = 0$ e $d = (0,0)$.*

Finalmente, estabelecemos o resultado fundamental desta seção.

Teorema 4.14 *O problema (P) é KT-invexo de segunda ordem se, e somente se, todo ponto Kuhn-Tucker de segunda ordem é minimizador global de (P).*

Demonstração. ($\Rightarrow$) Suponha que (P) é KT-invexo de segunda ordem e, por absurdo, suponha que existe um ponto Kuhn-Tucker de segunda ordem $x \in X$, e que x não é um minimizador local de (P). Então, existe $y \in X$ tal que $f(y) < f(x)$. Pela Proposição 4.11, (4.11) ocorre para toda direção crítica $d \in D(x)$. Mas, pela Proposição 4.10, existem $\eta(y,x) \in \mathbb{R}^n$, $\omega(y,x) \geqq 0$ e $d(x,y) \in D(x)$ tais que

$$\nabla^T f(x)\eta(y,x) + \omega(y,x)[d^T(y,x)\nabla^2 f(x) d(y,x)] = -1$$

$$\nabla^T g_i(x)\eta(y,x) + \omega(y,x)[d^T(y,x)\nabla^2 g_i(x) d(y,x)] \leqq 0, i \in I(x)$$

o que contraria a Proposição 4.11.

($\Leftarrow$) Agora, suponha que todo ponto Kuhn-Tucker de segunda ordem de (P) é um minimizador global de (P). Sejam $y, x \in X$ e $f(y) < f(x)$. Da hipótese feita, x não é um ponto Kuhn-Tucker de segunda ordem de (P). Pela Proposição 4.11, existe uma direção crítica $d \in D(x)$ que não cumpre a condição (4.11). Portanto, existe uma direção crítica $d(x, \overline{x})$, um número não negativo $\omega(x, \overline{x}) \geqq 0$ e uma direção $\eta(y, x) \in \mathbb{R}^n$, tal que,

$$\nabla^T f(x)\eta(y, x) + \omega(y, x)[d^T(y, x)\nabla^2 f(x)d(y, x)] < 0$$

e

$$\nabla^T g_i(x)\eta(y, x) + \omega(y, x)[d^T(y, x)\nabla^2 g_i(x)d(y, x)] \leqq 0, i \in I(x).$$

Mas, pela Proposição 4.10, isto é equivalente a dizer que (P) é KT-invexo de segunda ordem. $\square$

Parte II

Problema Multiobjetivo

Capítulo 5

Formulação do problema, conceitos de solução e escalarização

5.1 Formulação do problema e conceitos de solução

Na Segunda Parte deste trabalho, vamos considerar o problema de otimização multiobjetivo, o qual pode ser formulado como:

$$\begin{array}{l} \text{Minimizar } f(x) := (f_1(x), \cdots, f_p(x)) \\ \text{sujeito a:} \\ g_i(x) \leqq 0, \ i \in I \\ x \in U \end{array} \tag{MOP}$$

onde $U \subset \mathbb{R}^n$ é aberto e não vazio, e as funções $f_j, g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ são funções diferenciáveis em U.

Denotaremos: $J = \{1, \cdots, p\}$, $I = \{1, \cdots, m\}$.

O conjunto $X = \{x \in U : g_i(x) \leqq 0, i \in I\}$ é o conjunto factível de (MOP).

Cada ponto $x \in X$ é chamado um ponto factível de (MOP).

Para cada $x \in X$, definimos $I(x) = \{i \in I : g_i(x) = 0\}$ chamado conjunto dos índices de restrições ativas em x.

Vários conceitos de solução podem ser associados ao problema (MOP). Talvez, os mais conhecidos sejam os seguintes:

- Solução fracamente eficiente: $x^* \in X$ é *solução fracamente eficiente* (resp. *solução local fracamente eficiente*) de (MOP) se não existe $x \in X$ tal que $f(x) < f(x^*)$, (resp. se existe uma vizinhança N de x^* tal que não existe $x \cap N$ tal que $f(x) < f(x^*)$).

- Solução eficiente: $x^* \in X$ é *solução eficiente* (resp. *solução local eficiente*) de (MOP) se não existe $x \in X$ tal que $f(x) \leq f(x^*)$, (resp. se existe uma vizinhança N de x^* tal que não existe $x \cap N$ tal que $f(x) \leq f(x^*)$).

- Solução propriamente eficiente: $x^* \in X$ é *solução propriamente eficiente* de (MOP) se ela é eficiente e se existe $M > 0$ tal que, para cada $k \in J$

  $$\frac{f_k(x^*) - f_k(x)}{f_j(x) - f_j(x^*)} \leq M$$

  para algum j tal que $f_j(x^*) < f_j(x)$ quando $x \in X$ e $f_k(x) < f_k(x^*)$.

Observação 9 *Temos que:*

$$\begin{array}{l} \text{Maximizar } f(x) = (f_1(x), \ldots, f_p(x)) \\ \text{sujeito a:} \\ x \in X \end{array} \tag{5.1}$$

É equivalente à

$$\begin{array}{l} \text{Minimizar } -f(x) = -(f_1(x), \ldots, f_p(x)) \\ \text{sujeito a:} \\ x \in X \end{array} \tag{5.2}$$

É imediato das definições: eficiência própria $\Longrightarrow$ eficiência $\Longrightarrow$ eficiência fraca. As recíprocas são falsas, como mostram os seguintes exemplos:

Exemplo 5.1 *Moulin e Soulié [12].*
Seja o problema:

$$\begin{array}{l} \text{Maximizar } f(x) = (f_1(x), f_2(x)) \\ \text{sujeito a: } x \in X \end{array} \tag{5.3}$$

onde $f : \mathbb{R}^2 \to \mathbb{R}^2$ *com* $f_1(x) = x_1$ *e* $f_2(x) = x_2$, *e* $X = \left\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leqq 1\right\} \cup [-1, 0] \times [0, 1] \subseteq \mathbb{R}^2$.

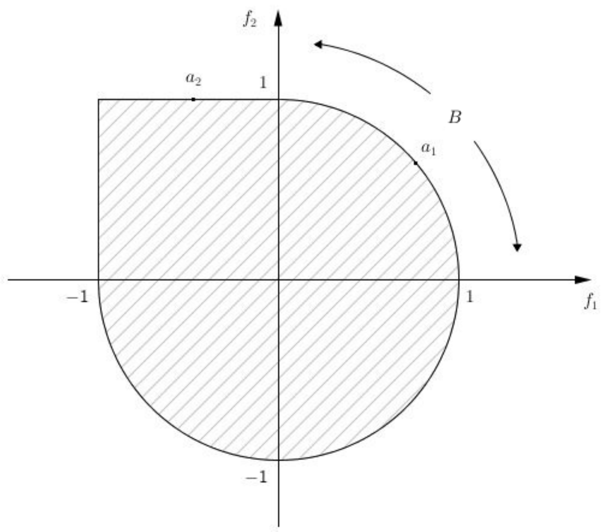


Figura 5.1: Conjunto factível X

- *O ponto* $a_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ *não pode melhorar estritamente* f_1, *sem diminuir estritamente* f_2. *O ponto* a_1 *é solução eficiente pertencente ao conjunto* B, *onde* B *é o arco compreendido entre* $(1, 0)$ *e* $(0, 1)$.

- *O ponto* $a_2 = (\frac{-1}{2}, 1)$ *pode melhorar* f_1 *sem diminuir estritamente* f_2, *(consideremos por exemplo, o ponto* $(0, 1)$*), mas não pode melhorar os dois critérios ao mesmo tempo. O ponto* $a_2 \in B_f$ *entretanto* a_2 *pode* $\notin B$. *Com:*

$$B_f = \{B \cup (\lambda, 1) : -1 \leqq \lambda \leqq 0\}$$

- *Na Figura 5.1 vizualizamos os pontos eficientes como sendo os pontos de $f(x)$, tais que, a interseção de $\alpha+\mathbb{R}^2_+$ com $f(x)$ se reduz aos pontos α e as soluções fracamente eficientes, como pontos β de $f(x)$, tais que, a interseção de $\beta+\left\{\mathbb{R}^2_+ \setminus 0\right\}$ com $f(x)$ é vazia.*

Exemplo 5.2 *Considere o problema:*

$$\begin{array}{l} \textit{Minimizar } f(x) = (f_1(x), f_2(x)) \\ \textit{sujeito a: } x \geqq 0,\ x \in \mathbb{R}. \end{array}$$

onde $f_1(x) = -x^2$ e $f_2(x) = x^3$. Temos que o ponto $x = 0$ é eficiente, porém não é propriamente eficiente. De fato, a razão

$$\frac{f_1(0) - f_1(x)}{f_2(x) - f_2(0)} = \frac{1}{x}$$

é ilimitada quando x tende a zero pela direita.

5.2 Escalarização

Por "escalarização" entende-se um processo através do qual é possível converter um problema multiobjetivo em um problema de otimização mono-objetivo (ou "escalar") equivalente, de tal maneira que as soluções do problema multiobjetivo podem ser obtidas como soluções de um problema clássico de programação não-linear.

Vamos considerar o *Método da Ponderação* e o *Método da ε-restrição.*

1. **Método da Ponderação**: Seja $W = \{w \in \mathbb{R}^p : w_j \geq 0, j \in J \text{ e } \sum_{j=1}^p w_j = 1\}$. Para cada $w \in W$, consideraremos o seguinte problema (ponderado)

$$\begin{array}{l} \text{Minimizar } \sum_{j=1}^p w_j f_j(x) \\ \text{sujeito a: } x \in X. \end{array} \qquad (\mathrm{P}_w)$$

2. **Método da ε-restrição:** Para cada vetor $\varepsilon \in \mathbb{R}^p$ e $k \in J$, definimos o problema

$$\begin{array}{l} \text{Minimizar } f_k(x) \\ \text{sujeito a:} \\ f_j(x) \leqq \varepsilon_j, j \in J,\ j \neq k \\ x \in X. \end{array} \qquad (\mathrm{P}_k(\varepsilon))$$

O objetivo desta Seção é descrever as relações existentes entre as soluções do problema multiobjetivo (MOP) e as soluções dos problemas escalarizados (P_w) e $(\mathrm{P}_k(\varepsilon))$.

Teorema 5.3 *Se existe $w \in W$ tal que $\overline{x} \in X$ é solução de (P_w), então $\overline{x}$ é solução fracamente eficiente de (MOP).*

Demonstração. Se $\overline{x}$ não fosse solução fracamente eficiente de (MOP), então existiria um $x \in X$ tal que $f_j(x) < f_j(\overline{x})$, para cada $j \in J$. Se multiplicamos estas desigualdades por $w_j \geq 0$ e somarmos sobre $j \in J$ obtemos $\sum_{j\in J} w_j f_j(x) < \sum_{j \in J} w_j f_j(\overline{x})$ o que é absurdo, pois $\overline{x} \in X$ é solução de (P_w). $\square$

Teorema 5.4 *Suponha que todas as funções de (MOP) sejam convexas. Se $\overline{x}$ é uma solução eficiente de (MOP) então existe $w \in W$ tal que $\overline{x}$ é solução de (P_w).*

Demonstração. Suponha que $\overline{x} \in X$ é solução eficiente de (MOP). Seja $x \in X$ e defina:

$$\begin{aligned} U &= \{u \in R^p : u > 0\} \\ V &= \{v \in R^p : v \leqq f(\overline{x}) - f(x)\}. \end{aligned}$$

Então, $U \cap V = \emptyset$. Além disso, U e V são convexos e não-vazios. Logo, pelo Teorema de Separação de Hahn-Banach (ver [11]) existe um vetor $w \in R^p$, $w \neq 0$ tal que

$$w^T v \leqq 0 < w^T u, \forall u \in U, \forall v \in V.$$

Portanto, $w \geqq 0$. É claro que $v = f(\overline{x}) - f(x) \in V$. Assim,

$$w^T \left(f(\overline{x}) - f(x)\right) \leqq 0$$

e, portanto

$$w^T f(\overline{x}) \leqq w^T f(x), \forall x \in X$$

ou, equivalentemente, $\overline{x}$ é solução de (P_w). □

Teorema 5.5 *Se existe $w \in W$ tal que $w > 0$, $\overline{x} \in X$, tal que, é solução de (P_w), então $\overline{x}$ é solução propriamente eficiente de (P).*

Demonstração. Vamos supor que $w > 0$ e que $\sum_{j\in J} w_j f_j(\overline{x}) \leqq \sum_{j\in J} w_j f_j(x), \forall x \in X$ e, por absurdo, suporemos que $\overline{x}$ não é solução eficiente de (MOP). Neste caso, existe $x_0 \in X$ tal que $f(x_0) \leq f(\overline{x})$, e como $w > 0$, teríamos $\sum_{j\in J} w_j f_j(\overline{x}) > \sum_{j\in J} w_j f_j(x_0)$, o que é absurdo.

Agora, definimos $M = (p-1) \max\limits_{j,k} \dfrac{w_j}{w_k}$. Vamos demonstrar que $\overline{x}$ é solução propriamente eficiente de (MOP), com respeito a este M.

Suponha por absurdo que $\overline{x}$ não é solução propriamente eficiente de (MOP) para este M. Então, para algum $k \in J$ e algum $x \in X$, temos

$$f_k(\overline{x}) - f_k(x) > M(f_j(x) - f_j(\overline{x}))$$

para cada j com $f_j(\overline{x}) < f_j(x)$ e $f_k(x) < f_k(\overline{x})$. Logo, para cada $j \neq k$ temos

$$f_k(\overline{x}) - f_k(x) > \frac{p-1}{w_k} w_j(f_j(x) - f_j(\overline{x}))$$

e, multiplicando a desigualdade acima por $\dfrac{w_k}{p-1}$ e somando todas as desigualdades obtidas sobre $j \neq k$, obtemos

$$w_k(f_k(\overline{x}) - f_k(x)) > \sum_{j\neq k} w_j(f_j(x) - f_j(\overline{x}))$$

donde

$$\sum_{j\in J} w_j f_j(\overline{x}) > \sum_{j\in J} w_j f_j(x)$$

o que contraria o fato de $\overline{x}$ ser solução ótima de (P_w). □

Em [25], Osuna-Gómez et al. introduzem o seguinte conceito de convexidade generalizada para o problema multiobjetivo (MOP):

Definição 5.6 *(Osuna-Gómez et al. [25]) Dizemos que o problema (MOP) é **KT-invexo** em X com respeito a η se, para quaisquer $x, y \in U$, com $g(x) \leqq 0, g(y) \leqq 0$, existe um vetor $\eta(y, x) \in \mathbb{R}^n$, tal que*

$$f_j(y) - f_j(x) \geqq \nabla^T f_j(x)\eta(y, x) \ \forall j \in J$$

$$-\nabla^T g_i(x)\eta(y, x) \geqq 0 \ \forall i \in I(x)$$

Antes de mostrarmos como os conceitos de KT-invexidade e de escalarização (ponderação) se relacionam, apresentamos dois lemas essenciais na prova do teorema que segue.

Lema 5.7 *Se o problema (MOP) é KT-invexo, então todo ponto Kuhn-Tucker é uma solução fracamente eficiente.*

Demonstração. Seja $\overline{x}$ um ponto Kuhn-Tucker para (MOP), e suponha que este problema é KT-invexo. Devemos concluir que $\overline{x}$ é um ponto fracamente eficiente.

Se, por absurdo, existe outro ponto factível $x \in X$ tal que $f_j(x) < f_j(\overline{x}), \forall j \in J$, então

$$0 > f_j(x) - f_j(\overline{x}) \geqq \nabla^T f_j(\overline{x})\eta(x, \overline{x}) \ \forall j \in J$$

o que implica,

$$\sum_{j \in J} \lambda_j \nabla^T f_j(\overline{x})\eta(x, \overline{x}) < 0, \ \forall \lambda_j \geq 0 \tag{5.4}$$

Como $\overline{x}$ é um ponto Kuhn-Tucker,

$$\sum_{j \in J} \lambda_j \nabla^T f_j(\overline{x})\eta(x, \overline{x}) + \sum_{i \in I(\overline{x})} \mu_i \nabla^T g_i(\overline{x})\eta(x, \overline{x}) = 0 \tag{5.5}$$

De (5.4), e (5.5), temos que

$$\sum_{i \in I(\overline{x})} \mu_i \nabla^T g_i(\overline{x})\eta(x, \overline{x}) > 0 \tag{5.6}$$

Dado que o problema (MOP) é KT-invexo, segue que

$$-\nabla^T g_i(\overline{x})\eta(x, \overline{x}) \geqq 0, \ \forall i \in I(\overline{x})$$

e desde que $\mu_i \geq 0$,

$$-\mu_i \nabla^T g_i(\overline{x})\eta(x, \overline{x}) \geqq 0, \ \forall i \in I(\overline{x})$$

Portanto,

$$\sum_{i \in I(\overline{x})} \mu_i \nabla g_i(\overline{x})\eta(x, \overline{x}) \leqq 0$$

Mas isto contradiz (5.6). □

Lema 5.8 *Se (MOP) é KT-invexo, então todo ponto Kuhn-Tucker resolve um problema escalar ponderado.*

Demonstração. Seja $\overline{x} \in X$ um ponto Kuhn-Tucker. Então, existe $\lambda \geq 0$ e $\mu \geq 0$ tal que

$$\begin{aligned} &\sum_{j\in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i\in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0 \\ &\mu_i g_i(\overline{x}) = 0, \forall i \in I(\overline{x}) \end{aligned} \tag{5.7}$$

Portanto $(\overline{x}, \mu)$ assegura as condições de Kuhn-Tucker para o problema escalar (P_λ). Além disso, se o problema (MOP) é KT-invexo, então o problema (P_λ) é KT-invexo e assim $\overline{x}$ é uma solução ótima de (P_λ) □

Com isso, temo o seguinte:

Teorema 5.9 *(MOP) é KT-invexo se, e somente se, para todo $\overline{x} \in X$ solução fracamente eficiente (MOP), existe $w \in W$ tal que $\overline{x} \in X$ é solução de (P_w)*

Demonstração. Nos Lemas (5.7) e (5.8) foi provada a suficiência do teorema. Mostraremos agora a recíproca.

Suponha que todo ponto Kuhn-Tucker é solução fracamente eficiente e resolve o problema escalar ponderado (P_w) para algum $w \in W$. Então, o seguinte sistema não tem solução, $(\lambda, v) \geq 0, v > 0$

$$\left\{ \begin{array}{l} \sum_{j\in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i\in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0 \\ v + \sum_{j\in J} w_j (f_j(x) - f_j(\overline{x})) = 0 \end{array} \right. \tag{5.8}$$

$\forall\, \overline{x}, x \in X$ com $g(\overline{x}) \leqq 0$ e $g(x) \leqq 0$. Ou equivalentemente,

$$\left\{ \begin{array}{l} \sum_{j\in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i\in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0 \\ \sum_{j\in J} w_j (f_j(x) - f_j(\overline{x})) = -v < 0 \end{array} \right. \tag{5.9}$$

$\forall\, \overline{x}, x \in X$ com $g(\overline{x}) \leqq 0$ e $g(x) \leqq 0$. Então, pelo Teorema de Motzkin (Lema 3.10), segue que o sistema abaixo tem solução (η, ε), $\eta \in \mathbb{R}^n$ e $\varepsilon \in \mathbb{R}$

$$\left\{ \begin{array}{l} \varepsilon < 0 \\ \nabla^T f_j(\overline{x})\eta + \varepsilon(f_j(x) - f_j(\overline{x})) < 0,\ \forall j \in J \\ \nabla^T g_i(\overline{x})\eta \leqq 0,\ i \in I(\overline{x}) \end{array} \right. \tag{5.10}$$

$\forall\, \overline{x}, x \in X$ com $g(\overline{x}) \leqq 0$ e $g(x) \leqq 0$. Logo, $\forall\, \overline{x}, x \in X$ com $g(x) \leqq 0$ e $g(\overline{x}) \leqq 0$, existe $\eta(x, \overline{x})$ tal que

$$\left\{ \begin{array}{l} f_j(x) - f_j(\overline{x}) \geqq \nabla^T f_j(\overline{x})\eta(x, \overline{x}),\ \forall j \in J \\ -\nabla^T g_i(\overline{x})\eta(x, \overline{x}) \geqq 0, i \in I(\overline{x}) \end{array} \right. \tag{5.11}$$

Ou seja, (MOP) é KT-invexo em X. □

Com hipóteses de convexidade, vale a recíproca do Teorema 5.5:

Teorema 5.10 *Suponha que o conjunto factível X de (MOP) é convexo e que as funções f_j, $j \in J$ são convexas em X. Então $\overline{x} \in X$ é solução propriamente eficiente de (MOP) se, e somente se, existe $w \in W$, $w > 0$ tal que $\overline{x}$ é solução de (P_w).*

Demonstração. ($\Leftarrow$) É válida sem quaisquer hipóteses de convexidade. Veja Teorema 5.5.

($\Rightarrow$) Se $\overline{x}$ é solução propriamente eficiente de (MOP), então existe $M > 0$ tal que para cada $k \in J$, o sistema

$$\begin{cases} f_k(x) < f_k(\overline{x}) \\ f_k(x) + Mf_j(x) < f_k(\overline{x}) + Mf_j(\overline{x}), \ j \neq k \end{cases}$$

não tem solução $x \in X$. Então pelo Teorema de Gordan Generalizado (Lema 3.2), para cada k fixado, existem λ_j^k, $j \in J$, com $\sum_{j \in J} \lambda_j^k = 1$ tal que $\lambda_j^k \geq 0$ e

$$\lambda_k^k f_k(x) + \sum_{j \neq k} \lambda_j^k (f_k(x) + Mf_j(x)) \geqq \lambda_k^k f_k(\overline{x}) + \sum_{j \neq k} \lambda_j^k (f_k(\overline{x}) + Mf_j(\overline{x})), \forall x \in X.$$

Agora, somando sobre $k = 1, \cdots, p$, obtemos

$$\sum_{j \in J} (1 + M \sum_{j \neq k} \lambda_j^k) f_j(x) \geqq \sum_{j \in J} (1 + M \sum_{j \neq k} \lambda_j^k) f_j(\overline{x}).$$

Basta tomar $w_j = (1 + M \sum_{j \neq k} \lambda_j^k)$. □

Os próximos resultados relacionam as soluções de (MOP) e as soluções de ($\mathrm{P}_k(\varepsilon)$).

Teorema 5.11 *(a) Se $\overline{x} \in X$ é solução eficiente de (MOP), então $\overline{x}$ é solução ótima de ($P_k(\varepsilon)$), para algum $k \in J$ e $\varepsilon \in \mathbb{R}^p$.*

(b) Seja $\overline{x} \in X$. Defina $\varepsilon_j = f_j(\overline{x})$, para cada $j \in J$. Se $\overline{x}$ é solução única de ($P_k(\varepsilon)$), então $\overline{x} \in X$ é solução eficiente de (MOP).

Demonstração.

(a) Se $\overline{x}$ é solução eficiente de (MOP), então não existe $x \in X$ tal que $f(x) \leq f(\overline{x})$. Definindo-se $\varepsilon_j = f_j(\overline{x})$, para cada $j \in J$, temos que $f_j(\overline{x}) \leqq \varepsilon_j$, para cada $j \in J$. Fixado $k \in J$ temos que $\overline{x}$ é factível para ($\mathrm{P}_k(\varepsilon)$). Além disto, $\overline{x}$ é solução ótima de ($\mathrm{P}_k(\varepsilon)$). (Caso contrário, existiria um $x \in X$ tal que $f_j(x) \leqq \varepsilon_j = f_j(\overline{x}), j \neq k$ e $f_k(x) < f_k(\overline{x})$, o que contraria a eficiência de $\overline{x}$.)

(b) Suponha que $\overline{x}$ é solução única de ($P_k(\varepsilon)$) e, por contradição, que $\overline{x}$ não é solução eficiente de (MOP). Então, existe $x \in X$ tal que $f_j(x) \leqq f_j(\overline{x})$ para cada $j \in J$, com uma desigualdade estrita válida para um certo $j_0 \in J$.

Então x é factível para ($P_k(\varepsilon)$) e $f_k(x) \leqq f_k(\overline{x})$. Logo, $f_k(x) = f_k(\overline{x})$, e, por unicidade, teríamos $x = \overline{x}$, o que é absurdo, pois $f_j(x) \leqq f_j(\overline{x})$ $\forall j \in J$.

□

Capítulo 6

Existência de soluções

O objetivo deste capítulo é garantir a existência de soluções para o problema (MOP).

Para isto, vamos considerar o problema de otimização multiobjetivo abstrato, isto é, cuja estrutura de ordenação dos vetores do espaço $\mathbb{R}^p$ é mais geral que a comparação de vetores $\leq$ e $<$, que utilizamos até aqui. Para tais problemas, consideraremos um resultado de existência de soluções, o qual será particularizado para o problema (MOP).

Definição 6.1 *Dizemos que uma relação binária $\prec$ em $Y \subset \mathbb{R}^p$ é uma **relação de dominação** (ou relação de preferência) se é transitiva e não-reflexiva. Dados $x, y \in \mathbb{R}^p$, se $x \succ y$ (ou, ainda $y \prec x$) diremos que x domina y.*

Definição 6.2 *Seja $\prec$ uma relação de dominação em Y. Diremos que $y^* \in Y$ é um **elemento eficiente** de Y se não existe $y \in Y$ tal que $y \prec y^*$. Denotamos $\mathcal{E}(Y, \prec)$ ao conjunto dos elementos eficientes de Y.*

Consideremos agora o problema (MOP) formulado anteriormente.

Com esta terminologia, teríamos $Y = f(X)$ e, a relação de dominação em Y será definida por:

$$y_1 = f(x_1) \prec y_2 = f(x_2) \iff y_1 \leq y_2.$$

Além disto, se $y^* = f(x^*)$ é elemento eficiente de $Y = f(X)$ então x^* é solução eficiente de (MOP). Neste caso, $\mathcal{E}(Y, \prec) = \{y \in \mathbb{R}^p : y = f(x^*),\ x^* \text{ é solução eficiente de (MOP)}\}$.

Definição 6.3 *Sejam $Y \subset \mathbb{R}^p$ e $\prec$ uma relação de dominação em Y. A multiaplicação*

$$\begin{array}{rcl} D & : & Y \rightrightarrows \mathbb{R}^p \\ y & \mapsto & D(y) = \{d \in \mathbb{R}^p : y \prec y + d\} \cup \{0\} \end{array}$$

*é chamada **estrutura de dominação** em Y.*

O conjunto dos elementos eficientes de Y, também pode ser denotado $\mathcal{E}(Y, D)$.

Retornemos ao problema multiobjetivo (MOP). Neste caso, a estrutura de dominação em $Y = f(X)$ é relação de dominação $\leq$, e dada por

$$D(y) = \{d \in \mathbb{R}^p : 0 \leq d\} \cup \{0\} = \mathbb{R}^p_+.$$

Recordemos que: $D \subset \mathbb{R}^p$ é um *cone* se, para cada $\lambda \in \mathbb{R}$, $\lambda \geq 0$ temos $\lambda D \subset D$. Se, além disto, tivermos $D + D \subset D$, o cone D é um *cone convexo*.

Um cone D é dito um *cone ponteado* se $D \cap (-D) = \{0\}$.

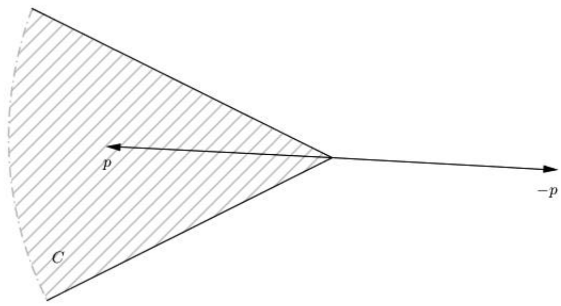


Figura 6.1: Cone Ponteado C

Particularmente interessantes são as estruturas de dominação em Y onde $D(\cdot) \equiv D$ (constante e igual a um cone em Y).

Observação 10 *Lembremos que a relação de ordem $\prec$ é assimétrica se, e somente se, $\forall\ x, y \in Y,\ x \succ y \Rightarrow y \not\succ x$.*

Proposição 6.4 *Suponha que a estrutura de dominação em Y é $D(\cdot) \equiv D$ (constante) e que D é um cone em Y e que $\prec$ é a relação de ordem associada. Então:*

(i) *$\prec$ é assimétrica se, e somente se, D é ponteado;*

(ii) *$\prec$ é transitiva se, e somente se, D é convexo.*

Demonstração.

(i) ($\Rightarrow$) Seja $\prec$ assimétrica e suponha, por contradição, que o cone D não é ponteado. Então, existe $z \neq 0$ tal que $z, (-z) \in D$. Assim, temos $0 \succ z$ e $0 \succ (-z)$. Sendo $\prec$ assimétrica $z \not\succ 0$ e $(-z) \not\succ 0$, implicando que $0 \notin z + D$ e $0 \notin (-z) + D$, ou seja, $z, (-z) \notin D$, o que é absurdo.

($\Leftarrow$) Agora, sejam D um cone ponteado e $x, y \in Y, x \neq y, x \succ y$. Então, $x \in y + D$ e, portanto, $x - y \in D$. Suponha por absurdo, que $y \succ x$. Então, $y \in x + D$, ou seja, $y - x = -(x - y) \in D$ e sendo D ponteado, $x - y = 0$, o que contradiz o fato de $x \neq y$.

(ii) ($\Rightarrow$) Seja $\prec$ transitiva e suponha por contradição, que D não é convexo. Então, existem $x, y \in D$ tais que $x + y \notin D$ (claro que $x \neq y$). Mas $x = (x + y) - y \in D$, isto é, $x + y \succ y$. Além disso, $y \in D$, então $y \in 0 + D$ implicando que $y \succ 0$. Logo pela transitividade de $\prec$, temos que $x + y \succ 0$. Portanto, $x + y \in D$ o que é absurdo.

($\Leftarrow$) Reciprocamente, sejam D convexo e $x, y, z \in Y$ tais que $x \succ y$ e $y \succ z$. Então, $x \in y + D$ e $y \in z + D$, ou seja, $(x - y), (y - z) \in D$. Sendo D convexo $(x - y) + (y - z) = x - z \in D$, isto é, $x \succ z$ e a relação é transitiva. $\square$

Definição 6.5 *Uma estrutura de dominação D é chamada **acíclica** se ela não possui ciclos, isto é, se para cada $m \in N$, não existem $y_1, \cdots, y_m \in Y$ tais que*

$$y_1 \prec y_2 \prec \cdots \prec y_m \prec y_1.$$

Retornemos ao estudo das soluções eficientes de (MOP). Neste caso, a relação de dominação em $\mathbb{R}^p$ é dada por $\leq$ e a correspondente estrutura de dominação é dada por $\mathbb{R}^p_+$, a qual é acíclica.

Um primeiro resultado de existência de soluções é dado a seguir:

Proposição 6.6 *Se a estrutura de dominação D é chamada* ***acíclica*** *e, se para cada $y \in Y$ os conjuntos $D(y)\backslash\{0\}$ são abertos e, se o conjunto Y é compacto e não vazio, então $\mathcal{E}(Y, \prec) \neq \emptyset$.*

Demonstração. Suponha por absurdo que $\mathcal{E}(Y, \prec) = \emptyset$. Então, para cada $y \in Y$ existe $\overline{y} \in Y$ tal que $\overline{y} \succ y$ ou equivalentemente

$$\forall y \in Y, \exists \overline{y} \in Y : y \in \overline{y} + D(\overline{y}) \diagdown \{0\}$$

ou seja

$$Y \subset \bigcup_{y \in Y} (y + D(y) \diagdown \{0\}).$$

Como por hipótese $D(y)\diagdown\{0\}$ são abertos e Y é compacto, então existem $y_1, \cdots, y_m \in Y$ tais que

$$Y \subset \bigcup_{i=1}^{m} (y_i + D(y_i) \diagdown \{0\}).$$

Em particular, para cada $j \in \{1, \cdots, m\}$, existe $i \in \{1, \cdots, m\}$ tal que $y_j \in (y_i + D(y_i)\diagdown\{0\})$, ou seja $y_i \succ y_j$, o que contraria o fato de D ser acíclica. Portanto, $\mathcal{E}(Y, \prec) \neq \emptyset$. □

Este resultado não é de grande utilidade na obtenção das soluções eficientes de (MOP), pois $D(y)\backslash\{0\} = \mathbb{R}^p_+\backslash\{0\}$ não é aberto.

Agora, estudemos a estrutura de dominação para as soluções fracamente eficientes de (MOP).

Neste caso, a relação é ordem em $Y = f(X) \subset \mathbb{R}^p$ é dada por

$$y_1 = f(x_1) \prec y_2 = f(x_2) \iff y_1 < y_2.$$

e neste caso, a estrutura de dominação é dada por $D = \text{int}\mathbb{R}^p_+ \cup \{0\}$. Logo, para cada $y \in Y$, os conjuntos $D(y) = \text{int}\mathbb{R}^p_+$ são, trivialmente abertos. $\mathcal{E}(Y, \prec) = \{y \in \mathbb{R}^p : y = f(x^*),\ x^*$ é solução fracamente eficiente de (MOP)$\}$.

Assim sendo, como consequência da Proposição 6.6, obtemos um resultado de existência de soluções fracamente eficientes de (MOP).

Corolário 6.7 *Suponha que em (MOP) se tenha f contínua e que o conjunto factível X é compacto. Então (MOP) possui soluções fracamente eficientes.*

Demonstração. Como f é contínua e X é compacto, então $Y = f(X)$ é compacto. O resultado desejado segue da Proposição 6.6. □

Na realidade, sob as mesmas hipóteses do Corolário 6.7, podemos garantir que (MOP) admite soluções eficientes.

Necessitaremos da seguinte generalização do conceito de compacidade:

Definição 6.8 *Sejam $D \subset \mathbb{R}^p$ um cone e $Y \subset \mathbb{R}^p$. O conjunto Y é chamado* ***D-semicompacto*** *se cada cobertura aberta de Y, da forma*

$$\{(y_i - \overline{D})^c : y_i \in Y,\ i = 1, \cdots, n\}$$

tem subcobertura finita.

(Estamos denotando por $\overline{A}$ o fecho de A e por A^c o conjunto complementar de A)

Proposição 6.9 *Seja D um cone convexo com $\overline{D}$ ponteado, Y não-vazio e D-semicompacto, então $\mathcal{E}(Y, D) \neq \emptyset$.*

Demonstração. Como $D \subset \overline{D}$, então $\mathcal{E}(Y, D) \supset \mathcal{E}(Y, \overline{D})$. Por isto, basta considerar o caso em que D é um cone convexo, fechado e ponteado.

Podemos então definir:

$$y_1, y_2 \in Y,\ y_1 \leq_D y_2 \iff y_2 - y_1 \in D.$$

Além disto, $y \in \mathcal{E}(Y, D)$ se, e somente se, y é minimal em $(Y, \leq_D)$. Vamos provar que Y é indutivo: Por absurdo, suponha que existe $\widehat{Y} = \{y_i : i \in I\} \subset Y$, totalmente ordenado e que não admite limitante inferior em Y. Então, não existe $y \in Y$ tal que $y \leq_D y_i, i \in I$. Ou seja, para todo $y \in Y$, existe $i \in I$ tal que $y \not\leq_D y_i$, ou equivalentemente, $\bigcap\limits_{i \in I}(y_i - D) \cap Y = \emptyset$. Logo, para cada $y \in Y$, existe $i \in I$ tal que $y \notin y_i - D$, isto é, $Y \subset \bigcup\limits_{i \in I}(y_i - D)^c = \bigcup\limits_{i \in I}(y_i - \overline{D})^c$. Portanto, sendo Y um conjunto D-semicompacto, existem $y_i \in \widehat{Y}$, $i = 1, \cdots, m$ tais que

$$Y \subset \bigcup_{i=1}^{m}(y_i - D)^c. \tag{6.1}$$

Por outro lado, sendo $\widehat{Y}$ totalmente ordenado, podemos supor $y_1 \leq_D y_2 \leq_D \cdots \leq_D y_n$ donde segue que

$$y_1 - D \subset \cdots \subset y_n - D$$

e, portanto

$$(y_1 - D)^c \supset \cdots \supset (y_n - D)^c. \tag{6.2}$$

Logo, segue de (6.1) e (6.2),

$$Y \subset (y_1 - D)^c. \tag{6.3}$$

Mas, $y_1 \in Y$ e $y_1 = y_1 + 0 \in y_1 + D$, o que contraria (6.3). Portanto, Y é indutivo e, pelo Lema de Zorn (veja [11]), $(Y, \leq_D)$ admite elemento minimal, o qual está em $\mathcal{E}(Y, D) \neq \emptyset$. Logo, $\mathcal{E}(Y, D) \neq \emptyset$. □

E, como consequência imediata da Proposição (6.9), obtemos:

Corolário 6.10 *Suponha que em (MOP) se tenha f contínua e que o conjunto factível X é compacto. Então (MOP) possui soluções eficientes.*

Capítulo 7

Condições necessárias de otimalidade

Neste capítulo estenderemos as condições necessárias de otimalidade de primeira e de segunda ordem ao problema multiobjetivo (MOP).

Observamos que muitos dos resultados apresentados nesta seção são extensões bastante naturais das correspondentes, obtidas para o problema mono-objetivo (P), discutido na primeira parte desta dissertação. Assim, as demonstrações dos teoremas que são análogas não serão apresentadas com detalhes.

7.1 Condições de primeira ordem

As noções de ponto Kuhn-Tucker e de ponto Fritz-John, se estendem de maneira natural ao problema multiobjetivo (MOP):

Definição 7.1 *(i) Dizemos que $\overline{x} \in X$ é um **ponto Fritz-John** de (MOP) se existem $\lambda \in \mathbb{R}^p_+, \mu \in \mathbb{R}^m_+$ não todos nulos, tais que,*

$$\sum_{j=1}^{p} \lambda_j \nabla f_j(\overline{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\overline{x}) = 0 \tag{7.1}$$

$$\mu_i g_i(\overline{x}) = 0, \ i = 1, \cdots, m. \tag{7.2}$$

*(ii) O ponto $\overline{x} \in X$ é um **ponto Kuhn-Tucker** de (MOP) se existem $\lambda \in \mathbb{R}^p_+, \mu \in \mathbb{R}^m_+$, tais que, $\lambda \neq 0$ e se cumprem as condições (7.1) e (7.2).*

Teorema 7.2 *Se $\overline{x} \in X$ é solução local fracamente eficiente de (MOP), então $\overline{x}$ é ponto Fritz-John de (MOP). Além disto, se a qualificação de restrição da independência linear se verifica em $\overline{x}$, então $\overline{x}$ é ponto Kuhn-Tucker de (MOP).*

Demonstração. (Esboço da prova) Seja $\overline{x} \in X$ uma solução fracamente eficiente de (MOP). Então, o sistema

$$\begin{cases} \nabla^T f_j(\overline{x})u < 0, j \in J = \{1, \cdots, p\} \\ \nabla^T g_i(\overline{x})u < 0, i \in I(\overline{x}) \end{cases}$$

não tem solução $u \in \mathbb{R}^n$. (Caso contrário, teríamos que os pontos da forma $\overline{x} + tu \in X$, para todo $t > 0$ suficientemente pequeno e $f_j(\overline{x} + tu) < f_j(\overline{x}), \forall j \in J$, contrariando a eficiência local fraca de $\overline{x}$).

Aplicando o Teorema de Motzkin (Lema 3.10) ao sistema acima, existem $\lambda \in \mathbb{R}^p_+, \mu \in \mathbb{R}^m_+$ não todos nulos e tais que

$$\sum_{j\in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i\in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0.$$

O resultado desejado é obtido tomando $\mu_i = 0$, $i \in I\backslash\{0\}$.

Por contradição, suponha que (7.1) é válido para $\lambda = 0$. Isto contraria a qualificação de restrição. □

7.2 Condições de segunda ordem

Nesta seção, estenderemos as condições necessárias de segunda ordem obtidas na primeira parte para o problema (P) ao problema multiobjetivo (MOP).

Ao longo desta seção admitiremos que as funções do problema (MOP) são de classe C^2.

A noção de direção crítica se estende de maneira natural ao problema (MOP): Diremos que $d \in \mathbb{R}^n$ é uma direção crítica de (MOP) em $x \in X$ se

$$\begin{aligned} \nabla^T f_j(\overline{x})d &\leqq 0, j \in J \\ \nabla^T g_i(\overline{x})d &\leqq 0, i \in I(\overline{x}). \end{aligned}$$

Denotaremos $D(x)$ o conjunto das direções críticas em x.

Para cada $x \in X$ e $d \in D(x)$, definiremos os seguintes conjuntos de índices:

$$\begin{aligned} J_1(x,d) &= \{j \in J : \nabla^T f_j(x)d = 0\} \\ I_1(x,d) &= \{i \in I(x) : \nabla^T g_i(x)d = 0\}. \end{aligned}$$

Teorema 7.3 *Suponha que $\overline{x} \in X$ é solução local fracamente eficiente de (MOP). Então para cada direção $d \in D(\overline{x})$, existem vetores $\lambda \in \mathbb{R}^p, \mu \in \mathbb{R}^m$, com $\lambda \geqq 0, \mu \geqq 0$, não todos nulos, tais que,*

$$\mu_i g_i(\overline{x}) = 0, i \in I \tag{7.3}$$

$$\lambda_j \nabla^T f_j(\overline{x})d = 0, j \in J \tag{7.4}$$

$$\mu_i \nabla^T g_i(\overline{x})d = 0, i \in I(\overline{x}) \tag{7.5}$$

$$\sum_{j\in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i\in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0 \tag{7.6}$$

$$d^T \left(\sum_{j\in J_1(\overline{x},d)} \lambda_j \nabla^2 f_j(\overline{x}) + \sum_{i\in I_1(\overline{x},d)} \mu_i \nabla^2 g_i(\overline{x}) \right) d \geqq 0. \tag{7.7}$$

Demonstração. Esquema da demonstração:

(1) Seja $\overline{x} \in X$ uma solução local fracamente eficiente de (MOP). Então, para cada direção $d \in D(\overline{x})$ fixada, o sistema

$$\begin{cases} \nabla^T f_j(\overline{x})z + d^T \nabla^2 f_j(\overline{x})d < 0, j \in J_1(\overline{x},d) \\ \nabla^T g_i(\overline{x})z + d^T \nabla^2 g_i(\overline{x})d < 0, i \in I_1(\overline{x},d) \end{cases}$$

não tem solução $z \in \mathbb{R}^n$. (A prova segue as mesmas linhas da demonstração do Teorema 3.11).

(2) Definimos $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ e $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, onde:

(i) A_1 é a matriz cujas linhas são os vetores $\nabla f_j(\overline{x})$, $j \in J_1(\overline{x}, d)$;

(ii) A_2 é a matriz cujas linhas são os vetores $\nabla g_i(\overline{x})$, $i \in I_1(\overline{x}, d)$;

(iii) b_1 é o vetor cujas componentes são $-d^T \nabla^2 f_j(\overline{x}) d$, $j \in J_1(\overline{x}, d)$;

(iv) b_2 é o vetor cujas componentes são $-d^T \nabla^2 g_i(\overline{x}) d, i \in I_1(\overline{x}, d)$;

(v) k é a soma das cardinalidades de $I_1(\overline{x}, d)$ e $J_1(\overline{x}, d)$.

Assim sendo se $\overline{x} \in X$ é uma solução local fracamente eficiente de (MOP), então o sistema $Az < b$ não tem solução $z \in \mathbb{R}^n$. Consequentemente, se definirmos o vetor $\widetilde{y} = (y, \cdots, y) \in \mathbb{R}^k$, então o problema linear

$$\begin{aligned} &\text{Maximizar } y \\ &\text{sujeito a : } Az + \widetilde{y} \leqq b \\ &y \in \mathbb{R}, z \in \mathbb{R}^k \end{aligned} \tag{P*}$$

tem valor ótimo $\overline{y} \leq 0$. O problema dual de (P*) é o seguinte problema

$$\begin{aligned} &\text{Minimizar } b^T \theta \\ &\text{sujeito a:} \\ &A^T \theta = 0 \\ &\textstyle\sum_i \theta_i = 1, \theta_i \geqq 0. \end{aligned} \tag{D*}$$

Pelo Teorema de Dualidade (Lema 4.12), o problema (D*) possui uma solução $\overline{\theta}$, satisfazendo $\overline{\theta} = (\overline{\lambda}, \overline{\mu})$, onde $\overline{\lambda} = (\overline{\lambda}_j)_{j \in J_1(\overline{x}, d)}, \overline{\mu} = (\overline{\mu}_i)_{i \in I_1(\overline{x}, d)}$. Basta definir:

$$\begin{aligned} \lambda_j &= \begin{cases} \overline{\lambda}_j, & \text{se } j \in J_1(\overline{x}, d) \\ 0\,, & \text{se } j \in J \setminus J_1(\overline{x}, d) \end{cases} \\ \mu_i &= \begin{cases} \overline{\mu}_i, & \text{se } i \in I_1(\overline{x}, d) \\ 0\,, & \text{se } i \in \{I(\overline{x}) \setminus I_1(\overline{x}, d)\} \cup \{I \setminus I(\overline{x})\}\,. \end{cases} \end{aligned}$$

□

Motivados pelo Teorema 3.11, Santos et al. [15] propuseram as seguintes noções de ponto Kuhn-Tucker de segunda ordem e de ponto Fritz-John de segunda ordem:

Definição 7.4 *Seja $\overline{x} \in X$ um ponto factível para (MOP).*

(i) Se para cada direção $d \in D(\overline{x})$, existem vetores $\lambda \in \mathbb{R}^p, \mu \in \mathbb{R}^m$, com $\lambda \geqq 0, \mu \geqq 0$, não todos nulos, tais que, se cumprem as condições (7.3)-(7.7), dizemos que $\overline{x}$ é um ***ponto Fritz-John de segunda ordem*** *de (MOP).*

(ii) Se para cada direção $d \in D(\overline{x})$, existem vetores $\lambda \in \mathbb{R}^p, \mu \in \mathbb{R}^m$, com $\lambda \geq 0, \mu \geqq 0$, tais que, se cumprem as condições (7.3)-(7.7), dizemos que $\overline{x}$ é um ***ponto Kuhn-Tucker de segunda ordem****.*

Corolário 7.5 *Seja $\overline{x} \in X$ solução local fracamente eficiente de (MOP) e suponha que o conjunto $\{\nabla g_i(\overline{x})\}_{i \in I(\overline{x})}$ é linearmente independente. Então, $\overline{x}$ é ponto Kuhn-Tucker de segunda ordem.*

A prova é consequência quase imediata do Teorema 3.11, e será omitida.

Capítulo 8

Condições suficientes de otimalidade

Neste capítulo, discutiremos algumas condições de convexidade generalizada para o problema (MOP), as quais garantem que um ponto Kuhn-Tucker (ou ponto Kuhn-Tucker de segunda ordem) é solução fracamente eficiente de (MOP). Os resultados deste capítulo generalizam os resultados do Capítulo 4 a problemas multiobjetivos.

8.1 Condições de primeira ordem

Inicialmente, observamos que sob hipóteses de invexidade, todo ponto Kuhn-Tucker é solução fracamente eficiente de (MOP). De fato:

Proposição 8.1 *Suponha que $\overline{x} \in X$ é um ponto Kuhn-Tucker do problema (MOP) e que as funções f_j, g_i ($j \in J$, $i \in I(\overline{x})$) são invexas com respeito a uma mesma η. Então $\overline{x}$ é solução fracamente eficiente de (MOP).*

Demonstração. Se $\overline{x}$ é um ponto Kuhn-Tucker de (MOP), então existem $\lambda \geq 0, \mu \geqq 0$ tais que

$$\sum_{j \in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i \in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0 \tag{8.1}$$

Logo, pelo Teorema de Motzkin (Lema 3.10), o sistema

$$\begin{aligned} \nabla^T f_j(\overline{x}) u < 0, \forall j \in J \\ \nabla^T g_i(\overline{x}) u \leqq 0, \forall i \in I(\overline{x}) \end{aligned} \tag{8.2}$$

não tem solução $u \in \mathbb{R}^n$.

Agora, tome $x \in X, x \neq \overline{x}$ e suponha por contradição que $f_j(x) < f_j(\overline{x}), \forall j \in J$. Como $x \in X$ temos que $g_i(x) \leqq 0 = g_i(\overline{x}), \forall i \in I(\overline{x})$.

Da invexidade de $f_j, \forall j \in J$ e $g_i, \forall i \in I(\overline{x})$ em $\overline{x}$ temos que $\forall x \in U$, existe $\eta(x, \overline{x})$ tal que

$$\begin{aligned} f_j(x) - f_j(\overline{x}) \geqq \nabla^T f_j(\overline{x}) \eta(x, \overline{x}) \\ g_i(x) - g_i(\overline{x}) \geqq \nabla^T g_i(\overline{x}) \eta(x, \overline{x}) \end{aligned}$$

e tomando $u = \eta(x, \overline{x})$ temos o absurdo, o que conclui a prova. $\square$

De maneira análoga ao que fizemos no Capítulo 4, consideraremos a classe mais ampla de problemas para a qual as condições de Kuhn-Tucker são necessárias e suficientes para a eficiência fraca.

Definição 8.2 *(Osuna-Gómez et al. [25]) Dizemos que o problema (MOP) é* ***KT-pseudoinvexo*** *se para quaisquer $x, y \in U$, com $g(x) \leqq 0$ e $g(y) \leqq 0$, existe um vetor $\eta(x,y) \in \mathbb{R}^n$ tal que*

$$f_j(y) < f_j(x) \implies \nabla^T f_j(x)\eta(x,y) < 0,\ j \in J$$

$$-\nabla^T g_i(x)\eta(x,y) \geqq 0,\ i \in I(x).$$

Teorema 8.3 *O problema (MOP) é KT-pseudoinvexo se, e somente se, todo ponto Kuhn-Tucker é solução fracamente eficiente de (MOP).*

Demonstração. ($\Leftarrow$) Suponha que todo ponto Kuhn-Tucker é uma solução fracamente eficiente de (MOP). Sejam $x, \overline{x} \in X$, com $f_j(x) < f_j(\overline{x}), \forall j \in J$. Então, $\overline{x}$ não é solução fracamente eficiente de (MOP) e, por hipótese, $\overline{x}$ não é ponto Kuhn-Tucker. Então, não existe $(\lambda, \mu) \geq 0$, $\lambda \neq 0$, tal que

$$\sum_{j \in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i \in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0.$$

Desse modo, pelo Teorema de Motzkin (Lema 3.10), o sistema

$$\begin{cases} \nabla^T f_j(\overline{x})d < 0,\ j \in J \\ \nabla^T g_i(\overline{x})d \leqq 0,\ i \in I(\overline{x}) \end{cases}$$

tem solução $d \in \mathbb{R}^n$. Definindo-se $\eta(x, \overline{x}) = d$, obtemos o resultado desejado.

($\Rightarrow$) Suponha que o problema (MOP) é KT-peseudoinvexo e seja $\overline{x} \in X$ um ponto Kuhn-Tucker de (MOP). Provaremos que $\overline{x}$ é solução fracamente eficiente de (MOP). Por absurdo, suponha que exista $x \in X$ tal que $f(x) < f(\overline{x})$. Como (MOP) é KT-pseudoinvexo, $\nabla^T f_j(\overline{x})\eta(\overline{x}, x) < 0$, $j \in J$, portanto,

$$\sum_{j \in J} \lambda_j \nabla^T f_j(\overline{x})\eta(\overline{x}, x) < 0, \forall \lambda \geq 0.$$

Por outro lado, como $\overline{x} \in X$ um ponto Kuhn-Tucker de (MOP), existem $\lambda \geq 0, \mu \geqq 0$ tais que

$$\sum_{j \in J} \lambda_j \nabla^T f_j(\overline{x}) + \sum_{i \in I(\overline{x})} \mu_i \nabla^T g_i(\overline{x}) = 0.$$

Em particular, teríamos

$$\sum_{i \in I(\overline{x})} \mu_i \nabla^T g_i(\overline{x})\eta(\overline{x}, x) = -\sum_{j \in J} \lambda_j \nabla^T f_j(\overline{x})\eta(\overline{x}, x) > 0. \tag{8.3}$$

Pela KT-pseudoinvexidade, $-\nabla^T g_i(x)\eta(x,y) \geq 0$, $i \in I(x)$, e como $\mu \geqq 0$

$$\sum_{i \in I(\overline{x})} \mu_i \nabla^T g_i(\overline{x})\eta(\overline{x}, x) \leqq 0$$

o que contraria (8.3). □

8.2 Condições de segunda ordem

De maneira análoga ao Capítulo 4, estudaremos condições de convexidade generalizada, de modo que caracterizemos a mais ampla classe de problemas para os quais as condições de Kuhn-Tucker de segunda ordem sejam necessárias e suficientes para a eficiência fraca. Além disto, também obteremos condições suficientes de segunda ordem para a eficiência.

A noção de KT-invexidade de segunda ordem introduzida por Ivanov [28] para problemas escalares (mono-objetivos) se estende de maneira natural a problemas multiobjetivos:

Definição 8.4 *(Santos et al. [15]) Dizemos que o problema (MOP) é **KT-pseudoinvexo de segunda ordem** se, para quaisquer $x, y \in X$, com $f(y) < f(x)$, existe uma direção $d(x, y) \in D(x)$, um número não negativo $\omega(x, y)$ e um vetor $\eta(x, y) \in \mathbb{R}^n$ tais que*

$$\nabla^T f_j(x)\eta(x,y) + \omega(x,y)d^T(x,y)\nabla^2 f_j(x)d(x,y) < 0, \ j \in J$$

$$\nabla^T g_i(x)\eta(x,y) + \omega(x,y)d^T(x,y)\nabla^2 g_i(x)d(x,y) \leqq 0, \ i \in I(x).$$

Enunciaremos agora um Teorema da Alternativa, utilizado na prova do próximo Teorema.

Lema 8.5 *[Teorema de Motzkin não homogêneo] Sejam $a_j, b_i, c_i \in \mathbb{R}^n$ e $\alpha_j, \beta_i, \gamma_i \in \mathbb{R}$ dados, com J, I^-, I^0 conjuntos finitos de índices e $J \neq \emptyset$. Então, o sistema (com varável $x \in \mathbb{R}^n$)*

$$\begin{cases} a_j x + \alpha_j < 0, \ j \in J \\ b_i x + \beta_i \leq 0, \ i \in I^- \\ c_i x + \gamma_i = 0, \ i \in I^0 \end{cases}$$

é impossível se, e somente se, o seguinte sistema (com variáveis $\theta_j, \lambda_i, \mu_i \in \mathbb{R}$)

$$\begin{cases} \sum_{j\in J} \theta_j a_j + \sum_{i\in I^-} \lambda_i b_i + \sum_{i \in I^0} \mu_i c_i = 0 \\ \sum_{j\in J} \theta_j \alpha_j + \sum_{i\in I^-} \lambda_i \beta_i + \sum_{i \in I^0} \mu_i \gamma_i = \theta_0 \\ \theta_0 + \sum_{j \in J} \theta_j > 0 \\ \theta_j \geq 0, i \in J \cup \{0\}, \lambda_i \geq 0, i \in I^- \end{cases}$$

é possível.

Para maiores detalhes veja [10].

Teorema 8.6 *O problema (MOP) é KT-pseudoinvexo de segunda ordem se, e somente se, todo ponto Kuhn-Tucker de segunda ordem é solução fracamente eficiente de (MOP).*

Demonstração. ($\Rightarrow$) Suponha por contradição que $\overline{x} \in X$ é um ponto Kuhn-Tucker de segunda ordem para (MOP) e que não é solução fracamente eficiente. Então, existe $y \in X$ tal que $f(y) < f(\overline{x})$, e como (MOP) é KT-pesudoinvexo de segunda ordem, existem $d(\overline{x}, y) \in D(\overline{x})$, $\omega(\overline{x}, y) \in \mathbb{R}_+$ e $\eta(\overline{x}, y) \in \mathbb{R}^n$ tais que

$$\nabla f_j(\overline{x})\eta(\overline{x},y) + \omega(\overline{x},y)d^T(\overline{x},y)\nabla^2 f_j(\overline{x})d(\overline{x},y) < 0, \ j \in J \tag{8.4}$$

$$\nabla g_i(\overline{x})\eta(\overline{x},y) + \omega(\overline{x},y)d^T(\overline{x},y)\nabla^2 g_i(\overline{x})d(\overline{x},y) \leq 0,\ i \in I(\overline{x}) \tag{8.5}$$

Seja (λ, μ), o par de multiplicadores associados à direção crítica $d(\overline{x}, y)$. Se multiplicamos (8.4) por $\lambda_j \geqq 0$ e (8.5) por $\mu_i \geqq 0$ e, somarmos as desigualdades obtidas, temos

$$\begin{array}{c} \sum_{j\in J} \lambda_j[\nabla f_j(\overline{x})\eta(\overline{x},y) + \omega(\overline{x},y)d^T(\overline{x},y)\nabla^2 f_j(\overline{x})d(\overline{x},y)] \\ + \sum_{i\in I(\overline{x})} \mu_i[\nabla g_i(\overline{x})\eta(\overline{x},y) + \omega(\overline{x},y)d^T(\overline{x},y)\nabla^2 g_i(\overline{x})d(\overline{x},y)] < 0. \end{array}$$

Rearranjando os termos,

$$\begin{aligned} 0 &= \sum_{j\in J} \lambda_j \nabla^T f_j(\overline{x}) + \sum_{i\in I} \mu_i \nabla^T g_i(\overline{x}) \\ &< -\omega(\overline{x},y)d^T(\overline{x},y) \left[\sum_{j\in J} \lambda_j \nabla^2 f_j(\overline{x}) + \sum_{i\in I(\overline{x})} \mu_i \nabla^2 g_i(\overline{x}) \right] d(\overline{x},y) \leqq 0 \end{aligned} \tag{8.6}$$

o que é absurdo.

($\Leftarrow$) Suponha que todo ponto Kuhn-Tucker de segunda ordem é solução fracamente eficiente de (MOP). Fixe $\overline{x} \in X$ e uma direção crítica $d \in D(\overline{x})$. Agora, considere os seguintes sistemas:

$$\left.\begin{array}{c} \nabla^T f_j(\overline{x})\eta + \omega d^T \nabla^2 f_j(\overline{x})d < 0,\ j \in J \\ \nabla^T g_i(\overline{x})\eta + \omega d^T \nabla^2 g_i(\overline{x})d \leqq 0,\ i \in I(\overline{x}) \end{array}\right\} \tag{S$_1$}$$

onde $\eta \in \mathbb{R}^n$ e $\omega \in \mathbb{R}_+$ e o sistema

$$\left.\begin{array}{c} f_j(y) - f_j(\overline{x}) < 0,\ j \in J \\ g_i(y) \leqq 0,\ i \in I(\overline{x}) \end{array}\right\} \tag{S$_2$}$$

onde $y \in U$.

Pelo Teorema de Motzkin não homogêneo (Lema 8.5), $\overline{x} \in X$ é um ponto Kuhn-Tucker de segunda ordem se, e somente se, (S_1) não tem solução (η, ω). Por outro lado, $\overline{x} \in X$ é solução fracamente eficiente de (MOP) se, e somente se, o sistema (S_2) não tem solução y. Como por hipótese, todo ponto Kuhn-Tucker de segunda ordem é solução fracamente eficiente de (MOP), os sistemas (S_1) e (S_2) são equivalentes. Portanto, (MOP) é KT-pseudoinvexo de segunda ordem. □

No que segue, obtemos condições suficientes de segunda ordem para a eficiência no problema (MOP).

A próxima proposição estabelece condições suficientes para a eficiência local.

Proposição 8.7 *Seja $\overline{x} \in X$ um ponto factível para (MOP) e suponha que, para cada $d \in D(\overline{x}) \setminus \{0\}$, existam vetores $\lambda \in \mathbb{R}^p_+, \mu \in \mathbb{R}^m_+$, não todos nulos, tais que,*

$$\sum_{j\in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i\in I} \mu_i \nabla g_i(\overline{x}) = 0 \tag{8.7}$$

$$\mu_i g_i(\overline{x}) = 0,\ i \in I \tag{8.8}$$

$$d^T \left(\sum_{j\in J} \mu_j \nabla^2 f_j(\overline{x}) + \sum_{i\in I} \lambda_i \nabla^2 g_i(\overline{x}) \right) d > 0 \tag{8.9}$$

então $\overline{x}$ é solução local eficiente de (MOP).

Demonstração. Suponha por absurdo que $\overline{x} \in X$ satisfaz as condições (8.7)-(8.9), mas não é uma solução eficiente local de (MOP). Então, existe uma sequência $(x_k) \subset X$ tal que $x_k \to \overline{x}$ e $f(x_k) \leq f(\overline{x})$, para todo k. Sem perda de generalidade, podemos supor que $x_k = \overline{x} + t_k d_k$, com $t_k \to 0^+$, $d_k \in \mathbb{R}^n$ e $||d_k|| = 1$, para todo k. Novamente, sem perda de generalidade, podemos supor que $d_k \to d \in \mathbb{R}^n$. Neste caso, pela fórmula de Taylor, segue

$$\begin{aligned} 0 &\geqq f_j(\overline{x}+t_kd_k) - f_j(\overline{x}) = [f_j(\overline{x}+t_kd_k) - f_j(\overline{x}+t_kd)] + [f_j(\overline{x}+t_kd) - f_j(\overline{x})] \\ &= [f_j(\overline{x}+t_kd_k) - f_j(\overline{x}+t_kd)] + [t_k\nabla^T f_j(\overline{x})d + \frac{t_k^2}{2}d^T\nabla^2 f_j(\overline{x})d + r_j(t_k)] \end{aligned} \tag{8.10}$$

onde $\lim\limits_{k\to\infty} \frac{r_j(t_k)}{t_k^2} = 0$. Como f_j é contínua, temos

$$\lim_{k\to\infty}[f_j(\overline{x}+t_kd_k) - f_j(\overline{x}+t_kd)] = 0 \tag{8.11}$$

e de (8.10) e (8.11), segue que

$$t_k\nabla^T f_j(\overline{x})d + \frac{t_k^2}{2}d^T\nabla^2 f_j(\overline{x})d + r_j(t_k) \leqq 0$$

para todo k suficientemente grande e, como $t_k > 0$, obtemos

$$\nabla^T f_j(\overline{x})d + \frac{t_k}{2}d^T\nabla^2 f_j(\overline{x})d + \frac{r_j(t_k)}{t_k} \leqq 0 \tag{8.12}$$

e fazendo $k \to \infty$ nesta última equação, obtemos $\nabla^T f_j(\overline{x})d \leqq 0$, para cada $j \in J$.

Similarmente, para cada $i \in I(\overline{x})$, temos que

$$\nabla^T g_i(\overline{x})d + \frac{t_k}{2}d^T\nabla^2 g_i(\overline{x})d + \frac{R_i(t_k)}{t_k} \leqq 0 \tag{8.13}$$

onde $\lim\limits_{k\to\infty} \frac{R_i(t_k)}{t_k^2} = 0$ e, portanto, $\nabla^T g_i(\overline{x})d \leqq 0$, isto é, $d \in D(\overline{x}), d \neq 0$. Multiplicando (8.12) por $\lambda_j \geqq 0$ e (8.13) por $\mu_i \geqq 0$ e somando as desigualdades obtidas, temos

$$\begin{aligned} 0 &\geqq \underbrace{\left[d^T \sum_{j\in J}\lambda_j\nabla f_j(\overline{x}) + \sum_{i\in I(\overline{x})}\mu_i\nabla g_i(\overline{x})d\right]}_{=0} + \left[d^T(\sum_{j\in J}\mu_j\nabla^2 f_j(\overline{x}) + \sum_{i\in I}\lambda_i\nabla^2 g_i(\overline{x}))d\right] \\ &= d^T\left(\sum_{j\in J}\mu_j\nabla^2 f_j(\overline{x}) + \sum_{i\in I}\lambda_i\nabla^2 g_i(\overline{x})\right)d \end{aligned}$$

o que contradiz (8.9). □

Por fim, estabeleceremos uma condição suficiente de segunda ordem para a eficiência global em (MOP).

Recordemos que uma função diferenciável $\phi : K \subset \mathbb{R}^n \to \mathbb{R}$ é chamada **quase-invexa** se para cada par de vetores $x, y \in K$, com $x \neq y$ existe um vetor $\eta(x, y) \in \mathbb{R}^n$ tal que

$$\phi(x) \leqq \phi(y) \implies \nabla^T\phi(y)\eta(x, y) \leqq 0.$$

Note que se ϕ for continuamente diferenciável e $\eta(x, y) = x - y$, então ϕ será quase-convexa.

Para cada $\overline{x} \in X$, definiremos o conjunto

$$\widehat{D}(\overline{x}) = \{d \in D(\overline{x}) : \nabla^T f_j(\overline{x})d = 0, \text{ para algum } j_0 \in J\}.$$

Proposição 8.8 *Suponha que $\overline{x} \in X$ é um ponto factível de (MOP), as funções f_j, g_i ($j \in J, i \in I(\overline{x})$) são quase-invexas com respeito a uma mesma η e que $\eta(x, \overline{x}) \neq 0$, para todo $x \neq \overline{x}$. Se para cada direção $d \in \widehat{D}(\overline{x}) \diagdown \{0\}$, existirem vetores $\lambda \in \mathbb{R}^p_+ \diagdown \{0\}, \mu \in \mathbb{R}^m_+$, tais que,*

$$\sum_{j \in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i \in I} \mu_i \nabla g_i(\overline{x}) = 0$$

$$\mu_i g_i(\overline{x}) = 0, \ i \in I$$

$$d^T \left(\sum_{j \in J} \mu_j \nabla^2 f_j(\overline{x}) + \sum_{i \in I} \lambda_i \nabla^2 g_i(\overline{x}) \right) d > 0$$

então $\overline{x}$ é solução eficiente (global) do problema (MOP).

Demonstração. Suponhamos por contradição que $\overline{x}$ não é solução eficiente de (MOP). Então existe $x \in U$ tal que

$$f(x) \leq f(\overline{x})$$

$$g_i(x) \leqq 0, \ i \in I.$$

Em particular, temos que $f_j(x) \leqq f_j(\overline{x})$ para cada $j \in J$ e $g_i(x) \leqq g_i(\overline{x})$, para $i \in I(\overline{x})$. Como estas funções são quase-invexas com respeito a η, temos

$$\nabla^T f_j(\overline{x}) \eta(x, \overline{x}) \leqq 0, \ j \in J$$

$$\nabla^T g_i(\overline{x}) \eta(x, \overline{x}) \leqq 0, \ i \in I(\overline{x}).$$

Vamos considerar dois casos: (i) $\nabla^T f_j(\overline{x}) \eta(x, \overline{x}) < 0, \forall j \in J$; (ii) $\nabla^T f_j(\overline{x}) \eta(x, \overline{x}) = 0$, para algum $j_0 \in J$.

Inicialmente, suponha $\nabla^T f_j(\overline{x}) \eta(x, \overline{x}) < 0, \forall j \in J$. Neste caso, o sistema

$$\left. \begin{array}{l} \nabla^T f_j(\overline{x}) d < 0, \ j \in J \\ \nabla^T g_i(\overline{x}) d \leqq 0, \ i \in I(\overline{x}) \end{array} \right\}$$

tem uma solução $d \in R^n$. Pelo Teorema de Motzkin (Lema 3.10), não existem $\lambda \geq 0$, $\mu \geq 0$, tais que,

$$\sum_{j \in J} \lambda_j \nabla f_j(\overline{x}) + \sum_{i \in I(\overline{x})} \mu_i \nabla g_i(\overline{x}) = 0.$$

O que contradiz as hipóteses feitas.

Agora, suponhamos que exista um $j_0 \in J$ tal que $\nabla^T f_j(\overline{x}) \eta(x, \overline{x}) = 0$. Neste caso, $\eta(x, \overline{x}) \in \widehat{D}(\overline{x}) \backslash \{0\}$. Como f_{j_0} é quase-invexa, temos que

$$f_j(x) \leqq f_j(\overline{x}) \implies \nabla^T f_j(\overline{x}) \eta(x, \overline{x}) = \lim_{t \to 0^+} \frac{f_j(\overline{x} + t\eta(x, \overline{x})) - f_j(\overline{x})}{t} \leqq 0$$

para todo $j \in J$. Consequentemente, para cada $j \in J$ fixado, temos $f_j(\overline{x} + t\eta(x, \overline{x})) \leqq f_j(\overline{x})$ para todo $t > 0$ e suficientemente pequeno. Usando uma expansão de Taylor,

$$0 \geqq f_j(\overline{x} + t\eta(x, \overline{x})) - f_j(\overline{x}) = t\nabla^T f_j(\overline{x}) \eta(x, \overline{x}) + \frac{t^2}{2} \eta^T(x, \overline{x}) \nabla^2 f_j(\overline{x}) \eta(x, \overline{x}) + r_j(t),$$

onde $\lim_{t \to 0} \dfrac{r_j(t)}{t^2} = 0$. Obtemos,

$$\nabla^T f_j(\overline{x}) \eta(x, \overline{x}) + \frac{t}{2} \eta^T(x, \overline{x}) \nabla^2 f_j(\overline{x}) \eta(x, \overline{x}) + \frac{r_j(t)}{t} \leqq 0 \tag{8.14}$$

e, similarmente, para cada $i \in I(\overline{x})$ fixado,

$$\nabla^T g_i(\overline{x})\eta(x,\overline{x}) + \frac{t}{2}\eta^T(x,\overline{x})\nabla^2 g_i(\overline{x})\eta(x,\overline{x}) + \frac{R_i(t)}{t} \leqq 0 \tag{8.15}$$

onde $\lim\limits_{t\to 0} \dfrac{R_i(t)}{t^2} = 0$. Agora, multiplicamos (8.14) por $\lambda_j \geqq 0$ e (8.15) por $\mu_i \geqq 0$ e somando as desigualdades obtidas, temos que

$$\begin{aligned}
0 \;\geqq\; & \underbrace{\left[\sum_{j\in J}\lambda_j \nabla f_j(\overline{x}) + \sum_{i\in I(\overline{x})}\mu_i \nabla g_i(\overline{x})\right]^T}_{=0} \eta(x,\overline{x}) \\
& + \frac{t}{2}\left[\sum_{j\in J}\lambda_j \eta^T(x,\overline{x})\nabla^2 f_j(x,\overline{x})\eta(x,\overline{x}) + \sum_{i\in I(\overline{x})}\mu_i \eta^T(x,\overline{x})\nabla^2 g_i(x,\overline{x})\eta(x,\overline{x})\right] \\
= \;& \frac{t}{2}\left[\sum_{j\in J}\lambda_j \eta^T(x,\overline{x})\nabla^2 f_j(x,\overline{x})\eta(x,\overline{x}) + \sum_{i\in I(\overline{x})}\mu_i \eta^T(x,\overline{x})\nabla^2 g_i(x,\overline{x})\eta(x,\overline{x})\right]
\end{aligned}$$

para todo $t > 0$ suficientemente pequeno. Por isso,

$$\sum_{j\in J}\lambda_j \eta^T(x,\overline{x})\nabla^2 f_j(x,\overline{x})\eta(x,\overline{x}) + \sum_{i\in I(\overline{x})}\mu_i \eta^T(x,\overline{x})\nabla^2 g_i(x,\overline{x})\eta(x,\overline{x}) \leqq 0$$

o que é absurdo, pois $\eta(x,\overline{x}) \in \widehat{D}(\overline{x})\backslash\{0\}$. □

Capítulo 9

Conclusões e possibilidades de trabalho futuro

Nesta Dissertação foram considerados problemas de otimização com um e com vários objetivos. Foram estudadas a existência de soluções para tais problemas e condições necessárias e suficientes de otimalidade de primeira e de segunda ordem.

Existem na literatura especializada muitos conceitos de funções convexas generalizadas. Tais conceitos se relacionam à suficiência das condições de Kuhn-Tucker para a otimalidade. Nosso interesse foi o de considerar a classe de problemas para os quais as condições de Kuhn-Tucker sejam necessárias e suficientes para a otimalidade. Ou, dito de outro de modo, a maior classe de problemas para os quais as condições de Kuhn-Tucker sejam necessárias e suficientes para a otimalidade.

Estudamos também resultados semelhantes a estes, utilizando para isto as condições de segunda ordem.

Nos Capítulos 2, 3 e 4 abordamos o estudo do problema mono-objetivo. No Capítulo 2 tratamos da existência de solução para tais problemas. Baseados no muito conhecido Teorema de Weierstrass procuramos relaxar as hipóteses de continuidade da função objetivo e de compacidade do conjunto. No Capítulo 3 apresentamos os clássicos Teoremas de Fritz John e Kuhn-Tucker para as condições necessárias de primeira ordem, e um teorema relacionado às condições necessárias de segunda ordem. No Capítulo 4 procuramos garantir que as condições necessárias (de primeira e segunda ordem) estabelecidas anteriormente, sejam também suficientes para a otimalidade. Tais resultados foram obtidos sob hipóteses de convexidade e generalizados a classes mais amplas de funções. Neste contexto temos que a KT-invexidade (e a KT-ivnexidade de segunda ordem) nos garante que todo ponto de Kuhn-Tucker (ou ponto Kuhn-Tucker de segunda ordem) é minimizador global do Problema (P), ou seja, esta é a classe mais ampla de problemas para a qual as condições de Kuhn-Tucker sejam verificadas.

Os Capítulos 5, 6, 7 e 8 tratam do estudo do problema multiobjetivo, e uma análise similar obtida ao problema mono-objetivo foi realizada. No capítulo 5 apresentamos a caracterização do problema, os conceitos de soluções e alguns resultados de escalarização; observamos que a KT-invexidade para o problema (MOP) é que garante que toda solução fracamente eficiente resolve um problema escalar ponderado. No Capítulo 6 consideramos o problema de otimização multiobjetivo abstrato e por meio dele apresentamos resultados que garantem a existência de soluções para o Problema de estudo (MOP). No capítulo seguinte apresentamos resultados para as condições necessárias (de primeira e segunda ordem) de otimalidade do Problema Multiobjetivo (MOP) os quais, são generalizações naturais dos apresentados no Capítulo 3. Por fim no Capítulo 8, discutimos algumas condições de convexidade generalizada para o Problema (MOP) e observamos que a KT-pseudoinvexidade (ou a KT-pseudoinvexidade de segunda ordem) é a classe mais ampla a qual garante que, pontos Kuhn-Tucker (ou pontos Kuhn-Tucler de segunda ordem) são soluções fracamente eficientes de (MOP).

O prosseguimento natural deste trabalho seria o de tentar obter resultados análogos a estes, para diferentes tipos de problema. Por exemplo, o problema de otimização entre espaços abstratos, com estrutura de dominação definida por cones convexos. Um conceito de KT-invexidade para tais problemas, em contexto diferenciável, foi considerado por Santos [14] em sua Tese de Doutorado. Uma possibilidade de trabalho futuro seria estender este resultado, possivelmente, relaxando as hipóteses de diferenciabilidade das funções do problema. Outra possibilidade, seria a de propor uma noção análoga à de KT-invexidade de segunda ordem, utilizando as condições de otimalidade de segunda ordem.

Outra possibilidade é o estudo dos chamados problemas anormais. Tais problemas surgem quando o multiplicador associado à função objetivo é nulo. Neste caso, as regras de multiplicadores usuais são de pouca valia na busca de soluções ótimas. Para o caso diferenciável, Avakov [7] e Izmailov [2] consideram, respectivamente, os problemas com restrições de igualdade e de desigualdade e conseguem estender as regras usuais de multiplicadores usuais a alguns casos anormais. Entretanto, os resultados obtidos pelos autores exigem um maior grau de diferenciabilidade das funções envolvidas. Uma questão interessante ainda em aberto é: seria possível estender tais resultados para o problema não diferenciável ou, pelo menos, diminuindo o grau de diferenciabilidade das funções do problema? E seria possível obter uma noção similar à de KT-invexidade para tais problemas? Recentemente, Hernández-Jiménez et al. [3], propuseram uma noção de KT-invexidade para problemas anormais. Uma possibilidade seria a de se estender a noção de KT-invexidade de segunda ordem para tais problemas.

Uma interessante técnica empregada na obtenção de condições de otimalidade para problemas de programação não linear são os chamados "Métodos de Deformação". O método consiste, a grosso modo, em "deformar"um problema de otimização em um problema mais simples. Entre outras hipóteses, esta deformação deve preservar os pontos ótimos. Bobylev [20, 21] estudou vários problemas de deformação e, para isto, foram utilizadas as condições necessárias de primeira ordem. Seria possível adequar as técnicas de deformação empregando-se as condições de segunda ordem?

Outra possibilidade é o estudo de dualidade para problemas multiobjetivos. Para o problema de programação não linear, diferentes problemas duais podem ser definidos. Talvez, os mais populares sejam os de tipo Lagrangeano, Mond-Weir e de Wolfe. Para se obter as relações de dualidade entre os problemas primal e dual, se exige alguma noção de convexidade (ou convexidade generalizada) nos dados do problema. Novamente, existe uma infinidade de funções convexas generalizadas, utilizada na obtenção de tais resultados. Recentemente, Ivanov [29] propõe um novo tipo de dual e introduz o conceito de problemas WD-invexos. Tal classe de problemas tem a seguinte propriedade: um problema é WD-invexo se, e somente se, o par de problemas primal-dual satisfaz a propriedade da dualidade fraca. Cremos ser possível estabelecer resultados similares a este para o problema multiobjetivo. Seria possível estabelecer um resultado similar para o problema de otimização abstrato?

Referências Bibliográficas

[1] A. Izmailov e M. V. Solodov. *Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brasil, 2009.

[2] A. F. Izmailov. *Optimality conditions for degenerate extremum problems with inequality-type constraints*. U.S.R.R. Comp. Maths. Math. Phys., vol. 34, no. 6, 1994, pg 723-736.

[3] B. Hernández-Jiménez, M. Rojas-Medar, R. Osuna-Gómez and A. Beato-Moreno. *Generalized convexity in non-regular programming problems with inequality-type constraints*. Vol 352, no. 2, 2009, pg 604-613.

[4] D. H. Martin.*The Essence of Invexity*. Journal of Optimization Theory and Applications: vol 47, No.1, 1985, pg 65-76.

[5] D. W. Peterson. *A review of constraint qualifications infinite-dimensional spaces*. SIAM Review 15, 1973, pg 639-654.

[6] E. L. Lima. *Curso de Análise volume 2*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brasil, 1981.

[7] E. R. Avakov. *Extremum conditions for smooth problems with equality-type constraints*. U.S.R.R. Comp. Maths. Math. Phys., vol. 25, no. 3, 1985, pg 24-32.

[8] E. W. Karas e A. A. Ribeiro. *Um curso de Otimização*, Notas de aula, UFPR, Curitiba, Brasil, 2011.

[9] F.J. Gould and J.W. Tolle. *A necessary and sufficient qualification for constrained optimization*. SIAM J. Appl. Math. 20, 1971, pg 164-172.

[10] G. Bigi and M. Catellani. *Second order optimality conditions for diferenciable multiobjective problems*. RAIRO Oper. Res. 34, 1999, pg 441-426.

[11] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010

[12] H. Moulin et F. F. Soulié. *La convexité dans les mathématiques de la décision*. HERMANN, Paris, 1979.

[13] L. B. Contese.*Introduccion a La Optimizacion con Restricciones*. Universidad de Chile -Departamento de Matematicas y Ciencias de la Computacion, Casilla, 5272 - Correo 3 - Santiago, 1984.

[14] L. B. Santos. *Algumas contribuições em otimização multiobjetivo*. Tese de Doutorado. Departamento de Matemática Aplicada. Universidade Estadual de Campinas. 2004

[15] L. B. Santos, R. Osuna-Gómez, B. Hernández-Jiménez y M. A. Rojas-Medar.*Necessary and sufficient second order optimality conditions for multiobjective with C^1 data*. Nonlinear Analysis 85, 2013, pg 192-203.

[16] M. A. Hanson and B. Mond. *Further Generalizations of Convexity in Mathematical Programming* . Journal of Information anf Optimization Sciences, vol.3, 1982, pg 25-32.

[17] M. Castellani. *On constraint qualifications in nonlinear programming.* Nonlinear Analysis 69, 2008, pg 3249-3258.

[18] M. S. Bazaraa, J. J. Goode and C. M. Shetty. *Constraints qualifications revisited.* Management Science. Vol. 18, no. 9, 1972.

[19] M. S. Bazaraa, J. J. Jarvis and H. D. Sherali. *Linear programming network flows.* John Wiley & Sons, Inc., segunda edição, Estados Unidos da América, 1997.

[20] N. A. Bobylev. *"Deformation method of investigation of nonlinear programming problems I"* Automation and Remote Control, vol. 50, no. 7, Part I. 1989, pg 917-924.

[21] N. A. Bobylev. *"Deformation method of investigation of nonlinear programming problems II"* Automation and Remote Control, vol. 50, no. 8, Part II. 1989, pg 1018-1016.

[22] O. L. Mangasarian. *Nonlinear Programming* SIAM, 10, 1994.

[23] R. Andreani, G. Haeser, M. L. Schuverdt and P. J. S. Silva. *A relaxed constant positive linear dependence constraint qualification and applications.* Mathematical Programming, 135, 2012, 255-273.

[24] R. G. Eustáquio. *Condições de Otimalidade e de Qualificação para Problemas de Programação Não-Linear.* Dissertação de Mestrado, Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Curitiba, 2007.

[25] R. Osuna-Gómez, A. Beato-Moreno y A. Rufian-Lizana. *Generalized convexity in Multiobjective Programming.* Journal Math. Anal. Appl.233, 1999, pg 205-220.

[26] V. Chankong and Y. Y. Haimes. *Multiobjective Decision Making Theory and Methodology.* North-Holland, Amsterdam, 1983.

[27] V. I. Ivanov and I. Ginchev. *Second-order Optimality Conditions for Problems with C^1 Data.* Journal Mathematical Analysis and Applications. 340, 2008, pg 646-657.

[28] V. I. Ivanov. *Seconde-order Kuhn-Tucker Invex Constrained Problems.* Journal Glob. Optim. 50, 2011, pg 519-529.

[29] V. I. Ivanov. *"Duality in nonlinear programming".* Optimim. Letters. DOI 10.1007 s11590-012-0512-6, 2012.

[30] V. Pareto. *Cours d'economie politique.* Rouge, Lausanni, Suíça, 1896.

[31] V. Solodov. *Constraint qualifications* in Wiley Encyclopedia of Operations Research and Management Science, John Wiley & Sons, Inc., 2010.

[32] Z. Xu. *Constraint qualifications in a class of nondifferentiable mathematical programming problems*. Journal of Mathematical Analysis and Applications 302, 2005, pg. 282-290.